



PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/133258>

Please be advised that this information was generated on 2017-12-05 and may be subject to change.



tinbergen *institute*

Kahneman and Tversky and the Making of Behavioral Economics

Floris Heukelom

In the 1950s and 1960s, mathematical psychology and behavioral decision research arose in the unique context of the University of Michigan. These two psychological programs gave rise to Daniel Kahneman and Amos Tversky's famous research of the 1970s. In the early 1980s, Kahneman and Tversky's research was incorporated into financial economics, resulting in the new field of behavioral finance. In the 1990s, behavioral finance broadened its scope and became behavioral economics, a dominant new research program in economics. Despite its explorations in various directions and its claim of Herbert Simon's heritage in the 2000s, behavioral economics remains firmly attached to the framework as set out by Kahneman and Tversky.

Floris Heukelom holds a Master in Economics from the University of Amsterdam. In 2004, he started his PhD at the same university. Since 2007, he has been working as an Assistant Professor of Economics at the Radboud University Nijmegen.

Kahneman and Tversky and the Making of Behavioral Economics

Floris Heukelom



*Kahneman and Tversky and the Making of
Behavioral Economics*

ISBN 978 90 361 0125 7

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 455 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

*Kahneman and Tversky and the Making of
Behavioral Economics*

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie, in het openbaar te verdedigen
in de Agnietenkapel op vrijdag 29 mei 2009, te 10:00 uur

door

Floris Heukelom
geboren te Ten Boer

Promotiecommissie

Promotor: prof. dr. J.B. Davis

Co-promotor: dr. H.B.J.B. Maas

Overige leden: prof. dr. G. Gigerenzer
prof. dr. D.W. Hands
prof. dr. H. Maassen van den Brink
prof. dr. E-M. Sent
prof. dr. J.H. Sonnemans

Acknowledgements

I remember walking through the corridors of the University of Amsterdam some five years ago, thinking of which professor might want to supervise my Master thesis. Marcel Boumans, with whom I was taking a Master course in History and Methodology of Economics, drew my attention to a recently appointed professor, John Davis, an American. I entered John's office the next day and searched for the words I needed to express the enthusiastic, but vague idea I had in my mind. John listened patiently, smiled and then replied "Sure, go ahead." If that Master thesis, and subsequently this dissertation, have been brought to satisfying conclusions, then it is mainly due to the advice and comments I have received from John, but above all it is because of his unconditional support and his unwavering trust in the abilities I myself was not so sure I possessed. John, thank you.

The first time I met my future co-promotor Harro Maas, was during a third-year Bachelor course in the history of economics. Harro put the texts we had read on the table, looked at our small group of students, and said: "Can somebody explain what is going on in these texts? I don't understand." Over the years I have learned that apart from the fact that Harro almost always does know what is going on in a text, his question stemmed from what unfortunately is a rare trait in the Dutch educational system: the unconditional stride for quality. I would like to thank Harro for challenging me to reach my full potential.

In a direct extension of the support I received from John and Harro, this dissertation could not have been completed without the stimulating environment I found amongst my colleagues in the research group of History and Methodology of Economics at the University of Amsterdam: Geert Reuten, Edith Kuiper, Peter Rodenburg, Dirk Damsma, Murat Kotan, and all the others. I am particularly grateful to Marcel Boumans for providing his continuous advice and comments as I gradually grew to understand the history of measurement in twentieth-century science.

I would like to thank the Tinbergen Institute for greatly enhancing my understanding of economics. Its rigorous courses provided me with the in-depth knowledge I needed to find my own position, and ensured a stimulating intellectual environment in which this dissertation could steadily ripen.

I am grateful to the Adaptive Behavior and Cognition group at the Max Planck Institute for Human Development in Berlin for giving me the opportunity to spend six months in their midst in the Spring semester of 2006. In particular, I would like to

thank its director, Gerd Gigerenzer, for his always quick and apt comments on whichever piece of text I sent him, and for his unreserved support of my work.

Two years ago, the Department of Economics at the Radboud University Nijmegen offered me the possibility to advance my academic career and has provided me with the stable and stimulating environment in which this dissertation could be completed. A special thanks goes to Esther-Mirjam Sent for her trust in my then still to-be proven teaching abilities, and her advice and encouragement in all kinds of academic matters. A special thanks also goes to Michelle Mellion for editing the English of the entire dissertation manuscript. Roger, Remco, Ellen, Sjoerd, André, Ramzi, Ferdy, Aagje, Sandor, and Robbert gave me a warm welcome and became my new friends, which is so important when moving to a new city and starting work at a new university.

Finally there are those, friends and family, past and present, who I would like to thank simply for having been there: Tijmen, Naomi, Tiago, Bas, Frank, Julia, Eva, Joris, Astrid, Iris, Maria, Henny, Anne, Daan and Manon who made life all the richer; my family, Bart, Markus, Rita, and Bert, without whom I would never have been able to accomplish this feat; and most of all Marielle, whose smile softens the world's hard edges.

I dedicate the following pages to my mother, forever gone but never far away.

To my mother
Bineke Mebius

Table of contents

1. The disunified sciences of economics and psychology	1
1. Understanding the relation between economics and psychology	
2. The disunity of science	
3. Scope and limits of this thesis	
2. Measuring decisions and measurement as decision in postwar psychology	15
1. Psychology and measurement in Michigan	
2. Psychology at the University of Michigan and the Institute for Social Research	
3. Mathematical psychology	
4. Decision theory and behavioral decision research	
4.1 Decision theory	
4.2 Behavioral decision research	
5. “Measurement theory in psychology is behavior theory”	
6. Conclusion	
3. Tversky’s behavioral deviations and Kahneman’s cognitive mistakes	48
1. Kahneman and Tversky before Kahneman and Tversky	
2. Caught between <i>a priori</i> axioms and behavioral deviations: Tversky’s research in the 1960s	
2.1 Decision theory	
2.2 Measurement theory	
2.3 Behavioral decision research	
2.4 Situating Tversky’s experimental method	
2.5 The road not taken: Elimination by Aspects	
2.6 Caught between <i>a priori</i> axioms and behavioral deviations	
3. Kahneman’s cognitive mistakes	
3.1 From correlational to experimental psychology	
3.2 Kahneman’s research 1961-1971	
3.3 Kahneman’s cognitive errors	
4. Conclusion	
4. Heuristics and Biases for psychology and economics, Kahneman and Tversky in the 1970s	70
1. The Kahneman and Tversky collaboration	
2. Heuristics and Biases	
3. Kahneman’s case-based reasoning	
4. How to understand Heuristics and Biases	
4.1 The collaboration	
4.2 Kahneman and Tversky’s experiments	
4.3 Kahneman and Tversky versus Simon	

5. Prospect theory: Heuristics and Biases for economics
6. Economics and psychology
 - 6.1 Normative versus descriptive
 - 6.2 Prospect theory as unification of economics and psychology
7. Assessing Kahneman and Tversky

5. What to conclude from psychological experiments? 99
How Smith and Thaler incorporated behavioral deviations in economics

1. From psychology to economics
2. Corroboration and incorporation of psychology's behavioral deviations in Smith's experimental economics
3. Thaler's financial economic anomalies and the creation of behavioral finance
4. Distinguishing experimental economics from the rising star of behavioral finance
5. Economics, behavioral decision research, and Kahneman and Tversky

6. Building and defining behavioral economics 121

1. Who are we?
2. Building behavioral economics
 - 2.1 Intertemporal choice and the dual system approach
 - 2.2 Are preferences innate to human nature or do they emerge through interaction with the environment?
 - 2.3 Understanding behavioral economics in terms of rationality
 - 2.4 Economic policy as a form of paternalism
3. Defining behavioral economics
4. Behavioral economics in the 2000s: stable, defined, but slightly schizophrenic

7. Changing meanings: disunity in economics and psychology 149

References 154

Samenvatting in het Nederlands 171

1. The disunified sciences of economics and psychology

1. Understanding the relation between economics and psychology

The two-fold aim of this thesis is to understand Daniel Kahneman and Amos Tversky's research, and to understand how this research has altered economics in fundamental ways. Thus, this thesis is an exercise in postwar history of the relation between economics and psychology. The difficulty with this exercise, however, is that it is both unavoidable and deeply problematic. To regard the history of Kahneman and Tversky's work as an interaction between the scientific fields of economics and psychology is unavoidable because that is the language in which economists have described and understood it.¹ More generally, the division between the different disciplines is so deeply ingrained in twentieth-century Western social science – institutionally, conceptually, rhetorically, financially and so on – that it is virtually impossible to bypass. A historical analysis pertaining to the relation between psychology and economics in the twentieth century has in one way or another to use or deal with this division.

Yet, the distinction is problematic. A first problem is that the labels 'psychology,' 'economics,' 'psychologist,' and 'economist,' are not stable entities in postwar science. For instance, judged by received training, non-economists who have won the Nobel memorial prize in economics include Kahneman, Herbert Simon, and a whole range of physicists and engineers, including in-between cases such as Vernon Smith, who received a BA in electrical engineering and a MA and PhD in economics. Or consider Colin Camerer, currently one of the leading behavioral economists who holds a PhD in behavioral decision research. Moreover, the same is true for psychology. Foremost postwar mathematical psychologists such as R. Duncan Luce, Patrick Suppes and David Krantz, for instance, received degrees in engineering or mathematics before migrating to psychology.

In addition, these postwar scientists were labeled economist or psychologist flexibly and depending on the occasion. Depending on the situation, Simon called himself a political scientist, economist, psychologist and mathematician. Mathematician Leonard Savage has been claimed to be an important economist by economists and an important psychologist by psychologists. Even on the level of

¹ I refer to 'Kahneman and Tversky' throughout this dissertation. The order of their names bears no significance.

individual publications the standard divisions are problematic. Von Neumann and Morgenstern's *Theory of Games and Economic Behavior* (1944) has been described as a major contribution to their field by economists, psychologists, biologists, and mathematicians. Mathematical psychologists Krantz, Luce, Tversky, and Suppes conceived their three-volume *Foundations of Measurement* (1971, 1989, 1991) to extend the work of economist Gérard Debreu. However, at the same time they described it as a contribution to the empirical sciences in general, that is physics, economics, psychology and others, and thus as a contribution to the “methodology” of science. Hence, although it has been fundamentally ingrained in postwar science, the distinction between the different disciplines that scientists have employed has been anything but stable or clearly defined.

There is a second reason for the problematic nature of the division between psychology and economics. If there is one constant in postwar Western economics and psychology it has been the attempt to cross the alleged boundary between the two disciplines and to make this boundary disappear. Letters and minutes in Luce's archive in Harvard University show that ever since the Ogburn Report was initiated by President Herbert Hoover in 1929, the attempt to unify the behavioral and social sciences has been a constant theme in the National Science Foundation's (NSF) recurring reports from committees on Basic Research in the Behavioral and Social Sciences.² In the late 1950s Ward Edwards created behavioral decision research (BDR), a new field in psychology that applied economic theories to psychological problems. Three decades later Kahneman and Tversky introduced an adjusted Edwards program back into economics. Simon for the better part of his career tried to use the insights he gained originally in political science to alter economic theorizing, which led him to produce a new theory in psychology. Vernon Smith developed his new experimental methods for economists in the 1950s in collaboration with psychologist Sidney Siegel, effectively applying a psychological method to economic questions. Fifty years later he defended his program as good economics against increasing criticism from behavioral economists by claiming heritage to the work of

² In 1929, President Hoover commissioned a committee of social scientists to report on trends in the social and behavioral sciences, an effort to augment the knowledge base for his social policy. *Recent Social Trends in the United States* was published in 1933 and regularly updated during the following decades [Smelser (1986), p.21]. In the 1984 edition, the members of the committee, senior historians, sociologists, economists, and psychologists, discussed how the integration of the different social and behavioral sciences could be institutionalized, given that in practice they were already closely related and largely overlapped [Based on minutes, letters and reports in Luce's archive at Harvard University].

political scientist and psychologist Simon. The 1952 Santa Monica conference organized by mathematician Robert Thrall and psychologist Clyde Coombs is often cited as a major event in the history of game theory in economics, in the history of mathematical psychology, and in the history of experimental economics. As much as the division into different disciplines has been part of postwar science, so has the crossing and dissolving of the boundary been a constant.

A third and more subtle problem is that in the postwar period economists and psychologists have understood themselves, each other, and the boundary separating them in different ways. By and large, economists have understood economics to be a positive science refraining from normative claims, leaving those to the policy makers. With respect to psychology, postwar economists when pressed have made a distinction between psychological assumptions and the scientific field of psychology. The received view was that economists made psychological assumptions to build their theories and models on. However, that did not mean that economics and psychology were part of the same scientific discipline. Psychology, in this view, was a science of human behavior independent of economics, although it could always be used as source to improve the psychological assumptions made by economists. For economists, the economics-psychological border lay somewhere in between the scientific field of psychology and the psychological assumptions made by economists.

Psychologists for their part understood psychology as a science of human behavior in which the label psychology was used as a broad concept covering a range of sometimes very different ways of understanding different parts of human behavior. Psychology was understood by psychologists as being basically a descriptive and explanatory science, but the normative or prescriptive aspect of it was never far away. In fields such as psychotherapy, organizational psychology, and behavioral decision research the descriptive and the normative parts were kept separate, but there was always a clear and direct link between the two. The psychologists who were concerned with economics understood economics as a sub-discipline of psychology. Psychology in their understanding referred to the total of all different forms of investigation into individual human behavior, whereas economics referred to the sub-field that investigates a single part of human behavior, namely economic behavior. Moreover, economics was understood by psychologists as a scientific field which focused on the normative aspects of economic behavior. Because psychologists considered both the descriptive and the normative to be legitimate parts of science,

this exclusive focus of economists on the normative theory was in itself not problematic. But what it did imply was that from the psychologists' perspective economists focused only on one aspect of their scientific project. Hence, for psychologists the border between economics and psychology lay somewhere between the normative theories of economic behavior and the remainder of the psychological investigation of human behavior.

A further difference between economics and psychology in their perception of themselves, each other and the border between them was the status of the question how the two were related. For economists living in the twentieth century the relation between economics and psychology was something that was always in the back of their minds. It was always an issue that had to be dealt with, one that could never be dealt with satisfactorily and hence a question that would always reappear. Sometimes an economist would take a particularly clear and authoritative stance in this regard which made the issue disappear for a time. But it was always there and it always returned. For psychologists, on the other hand, the relative understanding of both disciplines was self-evident and rarely discussed. One of the surprises for the historian of economics who dives into the history of psychology is the absence of the question regarding how psychology is related to economics. It is not that the question simply is ignored, for psychologists do ask themselves from time to time how their discipline relates to economics, sociology, political science and other sciences. But in contrast to the relation with sociology, the relation between economics and psychology has always been clear to psychologists.³ At the end of the twentieth and beginning of the twenty-first century, and after thirty years of active use of psychology by economists, psychologists could complain that although "economists are now becoming more psychologically receptive, it is unfortunately less apparent that psychology is becoming more economically receptive" [Murnighan and Ross (1999), p.7], and that "[u]ntil recently, economists were more active in using and referring to social psychology than social psychologists were in using economics" [De Cremer, Zeelenberg, and Murnighan (2006), p.7]. The relation between the fields of economics and psychology, in other words, has been very much a concern of the economists.

³ It is true that many psychologists from the late nineteenth century onwards challenged psychology's hedonistic/utilitarian basis [Lewin (1996), pp.1299-1300], which served as a departure point for economics also. But this discontent was directed at philosophy, not economics.

For the historian writing about the postwar history of economics and psychology, the last point is particularly problematic because it renders the existing literature on the relation between economics and psychology in the two disciplines problematic as a basis on which to conduct further research. The accounts of both practicing scientists and historians in each of the disciplines of the relation between the two disciplines are largely incompatible. References to and accounts of Louis Leon Thurstone's work are illustrative in this regard. In postwar economics Thurstone (1931), "The Indifference Function" has been a constant reference for economists and historians of economics discussing the relation between economics and psychology. In these accounts Thurstone was the first to experimentally test economic demand theory. His experiments were initially set aside as irrelevant by most economists and followed by only a few [Moscati (2007)]. With the emergence of psychological experimentation in economics in the final quarter of the twentieth century, however, Thurstone was rediscovered and became regarded as an important precursor to contemporary research. Thus, Thurstone (1931) plays a significant role in economists' thinking about the relation between psychology and economics: in the early postwar years it was dismissed as of possible use in economics, and in the last few decades it was an important precursor for ways in which psychology could inform and improve economics' psychological assumptions.

Although Thurstone (1931) plays a role in economics and in histories of economics, Thurstone is not considered a major figure by economists. In psychology, however, Thurstone is seen as a key figure. Thurstone greatly improved measurement methods for social psychological research, thus ensuring the scientific status of social psychology. Applying his own measurement theory, Thurstone furthermore initiated a program of attitude measurement that became a cornerstone of modern psychology. Moreover, in the process he played a key role in uniting the fiercely opposed experimental and correlational (or Pearsonian) psychology. But in all this recognized importance of Thurstone's work and the accompanying discussion of his major works, Thurstone (1931) is completely absent. Thurstone's one-shot attempt to use the social psychological method to test economists' demand theory does not play any role in psychology, either in the past or in the present. Thus, whereas Thurstone seems to economists, both practitioners and historians, as an illustration of the relation between twentieth century economics and psychology in fact this only illuminates how economists and their historians have conceived of the relation between economics and

psychology. On the other hand, psychologists and their historians who consider economics to be a subfield of psychology view Thurstone as a major figure, but not with regard to the question of how economics and psychology are related to one another.

The unavoidable but problematic division of parts of postwar science into psychology and economics and especially the very different understanding of themselves, the other, and the border between them presents a dilemma for the historian who wants to do justice and to draw on both sides to provide a coherent understanding of a particular episode in the relation between the two. To understand why a particular branch of psychology could come to influence economics from the early 1980s onwards, and especially to understand in what way the experimental results and theories of this psychology were adopted and adapted to the economic framework, we need to understand the history of economics and psychology independently of one another. That is, we need to dive into the history of economics independently of how psychology thought and thinks of economics, and in particular, we need to draw out some of the threads in the history of psychology independently of how economists conceived and conceive of this history. We need to understand what was important in economics according to economists, and what was important in psychology according to psychologists. But at the same time we need to set out a story that will show us how the two scientific disciplines understood each other, and how they came to interact. We do not want two independent histories of psychology and economics, but two historical accounts that can be related. This boils down to the historiographical question of how to understand different, but related disciplines.

2. The disunity of science

Galison and Stump (1996) and Galison (1999) use the notion of the disunity of science to capture the idea that sciences and scientific practices may be separate and different, but at the same time be communicating and mutually influence each other. Galison applies disunity to the case of twentieth century physics in which he distinguishes three subcultures: theorizing, experimenting and instrument making: “[d]ifferent finite traditions of theorizing, experimenting, instrument making, and engineering, meet – sometimes even transform one another – but for all that they do not lose their separate identities and practices” [Galison (1999), p.137]. He makes two central points. First, contrary to the logical positivist tradition in science, there is not

one unified basis upon which different scientific subcultures are conducted.

“Experimentalists – and one could make a similar statement about theorists and instrumentalists – do not march in lockstep with theory. For example, the practice of experimental physics in the quantum mechanical revolution of 1926-27 was not violently dislocated *despite* the startling realignment of theory” [Galison (1999), p.143]. Second, contrary to the anti-positivist tradition in the history and philosophy of science, neither are different sciences and scientific practices entirely unrelated [Galison (1999), p.143].

Galison makes a comparison with the interactions of geographically scattered cultures that are studied by the anthropologist and introduces the concept of the trading zone. Different (sub-)cultures influence each other, for instance in their language and cultural habits. They trade with one another and goods travel from one culture to the next. But despite all their trading and mutual influencing, they remain different cultures nevertheless. Moreover, it is not necessarily true that because of the frequent cultural exchanges and trade, the different cultures are brought closer to one another over time. A cultural habit that travels from one culture to the next may acquire an entirely different meaning. Traded goods may be used for an entirely different purpose in the culture that they end up in compared to the culture in which they originated.

Consider the following example (this example is derived from Taussig (1980), and is used by Galison (1999)). In the Cauco Valley in Columbia, two groups of people live among each other each with their own culture. One culture consists of black peasants, descendants from slaves, running shops or working on the vast sugarcane farms. The peasants maintain a culture with magical cycles, sorcery and curing. The other culture is that of the rich, white landowners. The cultures exist alongside each other and frequently interact, for instance when a member of the landowner culture exchanges money for some eggs with a member of the peasant culture. The two cultures are, in other words, perfectly able to communicate with one another in specific contexts and can even be said to depend upon each other for their survival. However, the understanding of the exchange of money for eggs may be entirely different for the members of the two cultures. For a member of the landowner class, money is a neutral means of exchange that can accumulate into capital. For a member of the peasant culture the bank note possesses animistic and moral properties. In the most telling instantiation of this aspect of peasant culture, the godparent-to-be

hides a peso in his hand when a newborn is baptized by the Catholic priest. By doing so, also the peso bill is baptized and obtains the child's name, and the godparent-to-be also becomes the godparent of the peso bill. When this peso bill is put into circulation it will always come back to its owner when it is silently called upon three times. As a result, the members of the two cultures may be comfortable exchanging eggs for money with one another, but in a broader sense they interpret this transaction in a completely different way. For the landowner it is a neutral exchange of money for eggs, whereas for the peasant it signifies the return of the baptized peso bill to its owner.

Galison urges us to think of exchanges between different scientific (sub-) cultures in a similar fashion. At some level the different cultures may devise a context or set of rules within which they can exchange ideas, experimental results and instruments. Each is fluent in the exchange process in the sense that a member of the other culture will behave exactly as anticipated. But in a broader cultural context, the members of the different cultures will interpret the exchange in a completely different manner. "[T]he trading partners can hammer out a *local* coordination despite vast *global* differences" [Galison (1999), p.138]. The two cultures may entirely disagree about the implications of the information exchanged or its epistemic status. But at the same time "there is a context within which there is a great deal of consensus" [Galison (1999), p.146]. Depending on the topic to which we apply this anthropological analogy, we could also think of this context as a pidgin language. The different cultures devise a language that allows for a smooth exchange between the two cultures, but through back-and-forth trial and error and compromising devise a language that will never be able to satisfactorily capture each culture individually.

I apply Galison's approach to the case of economics and psychology. Economics and psychology are disunified cultures. They exist alongside each other and exchange results, instruments and ideas. They find it relatively easy to talk to one another, especially those representatives of the two cultures who operate in each other's vicinity. Indeed, one might say that for those scientists that operate close to the boundary on each side, the boundary seems more of a gradual continuum than a sharp line drawn in the sand. Both are also affected by the same challenges of the larger world and may at times come up with responses and adjustments that are very much alike. Yet despite all this exchanging and sharing of results, instruments and ideas, i.e. despite all their local coordination, they remain two clearly distinguished and

distinguishable sciences. Economists and psychologists have a lot in common, and yet are very different. They encounter each other frequently, they draw partly on the same authoritative sources, they inspire and influence each other, and psychologists have won the Nobel memorial prize in economics twice. At the same time, however, it is clear to both psychologists and economists that the two have been different in the past, that they are different at present, and that they will be distinguishable academic disciplines in the foreseeable future.

The disunity approach neatly captures how economists and psychologists talk about their relationship. Murnighan and Ross (1999), for instance, argue that although closely related, there will always be an “(invisible) dividing line between microeconomics and social psychology” [Murnighan and Ross (1999), p.2] because “the two fields promote different kinds of thinking and different philosophies, and these differences make it difficult for people in the two disciplines to collaborate, much less appreciate each other’s work” [Murnighan and Ross (1999), p.6]. The difference is that “[t]he objective of much of social psychology is to better understand how individuals make decisions in social situations. [...] Economics, on the other hand, is ultimately about explaining aggregates like market prices and quantities, incomes, employment and market efficiency” [Murnighan and Ross (1999), p.3]. Another good example is how one of these psychologists, Keith Murnighan, and an economist, Alvin Roth, recall the beginning of their rare thirty-years psychology-economics collaboration:

In essence, I did not speak economics and Al [Roth] did not speak psychology. (It is still not clear whether either of us has really picked up the other language, but at least we now think that we understand each other.) We had to work our way through a lengthy process to determine how we could express what we wanted to say about our joint work without offending each other or insulting each other’s fields in the process. [Murnighan and Roth (2006), pp. 322-323]

Murnighan and Roth, in other words, had to invent a pidgin language in order to be able to communicate. But that does not mean that they now, after thirty years of collaboration, understand the other’s native language.

The anthropological analogy also captures the idea that the different cultures may not be equally happy with the existing form of their interaction. In the interaction between economists and psychologists in the late twentieth and early twenty-first century it was the psychologists who sometimes felt abused in the exchange relation. A good example can be found in Lunt (1996). Lunt first of all notes that although economics and psychology may be investigating the same phenomena and relying on the same methods, “the way that social psychology approaches the study of the agent is very different from that of economics” [Lunt (1996), p.280]. But Lunt takes it a step further. It is not only that the two approach the agent in a different way, the exchange between economists and psychologists on this topic is also more favorable to economics, and in fact not at all to psychologists’ advantage. “[E]conomists work with simplified and anachronistic applications of psychological theory. We [psychologists] have to understand that psychology has become a resource for the economist, and [that] the motivation for integration is all on the side of psychology.” Furthermore, Lunt emphasizes that

[...] this is made particularly problematic by the kind of psychology utilized within economic theory. In my view, economists are not ready, prepared or even vaguely interested in changing their core assumptions as a response to psychological work. Indeed, we should realize that if an economist sounds interested in our work they are only trying us out to see what kind of resource we have to offer. The agenda for their interest will be some debate in economics that we won’t have even heard of. [Lunt (1996), p.283]

The different parties to the cultural exchange understand the exchange and the object exchanged in a different way. But on top of this, they may not be equally happy with the exchange, so that some members of one of the cultures may start to argue that they are being exploited.

3. Scope and limits of this thesis

The first two chapters of this thesis in particular are focused on the context in which scientific developments have taken place, yet this thesis is primarily a history of ideas. In addition, although the greater part of the history described in this thesis took place on American soil or was otherwise strongly connected to the United States, it is not

exclusively an American history. First, because the scientific developments described are a continuation of earlier science that has its origins outside the United States, principally in Europe. Second, because the two main protagonists of the present history, Kahneman and Tversky, although substantially Americanized, conducted the research for which they became famous at Israeli universities. Finally, the focus in this thesis will be on the developments leading to behavioral economics and on the development of behavioral economics itself. But I will use experimental economics as well to portray the differences between them.

The remainder of this thesis has been organized as follows. The upcoming second chapter discusses the work of the mathematical psychologists and behavioral decision researchers. I place this work in the unique context of the social and psychological research at the University of Michigan in the 1950s and 1960s. Although much attention has been paid to the Institute of Social Research (ISR), mathematical psychology and behavioral decision research at the University of Michigan have not yet been thoroughly investigated. However, these research programs and their related institutes would foster a development in psychology that was not only influential in the psychological discipline, but that would also transform economics from the 1980s onward. These research programs contributed to the development of both experimental economics, which developed from the 1960s onwards, and to Kahneman and Tversky-inspired behavioral economics that developed from the early 1980s onwards.

In the 1950s and 1960s the University of Michigan was arguably the center of American psychology, hosting the Institute for Social Research (ISR), Coombs' Michigan Mathematical Psychology Program, and Edwards' Engineering Psychology Laboratory and its related field of behavioral decision research. The second chapter argues that the key to understanding mathematical psychology and behavioral decision research is to see that, although largely separated and focused on different questions, both presumed the same two-sided understanding of psychology. In order to measure, one needed a sound theory of the measurement instrument, which was the human decision maker. Psychology at the same time measured human decision behavior and investigated the human being as a scientific measurement instrument.

This double understanding of psychology as using a measurement instrument to investigate that same measurement instrument became problematic when it turned out that the measurement instrument did not behave as it should. That was the

problem Tversky struggled with. Tversky had to choose between declaring the experimental results invalid and saying that the received theory of the measurement instrument was incorrect. Kahneman came to the rescue by suggesting that the human decision maker systematically and predictably deviates from how it should behave. Thus, the experimental results could be accepted, while at the same time the axioms of the measurement theory could be maintained. It did, however, give psychology the new task of investigating how and when human decision makers deviate from how they should behave. That new task was the basis of Kahneman and Tversky's collaborative research of the 1970s.

The third chapter, then, discusses the work Kahneman and Tversky did before their collaborative research. Tversky was educated at and received his PhD in the early 1960s from the University of Michigan under the supervision of Coombs and Edwards. Tversky's research embodied the synthesis of mathematical psychology and behavioral decision research. Towards the late 1960s, however, Tversky increasingly struggled with the tension between Leonard Savage's *a priori* axioms of decision theory and the behavioral deviations he observed in his experiments. Kahneman, for his part, came from a very different background. Strongly influenced by his experience as psychologist in the Israeli army, Kahneman's different research interests focused on human's cognitive mistakes. Kahneman showed that despite the fact that we think we do cognitively quite well in the course of our daily lives, in fact we constantly make systematic cognitive mistakes.

In 1969 Kahneman and Tversky started their long and fruitful collaboration. The most productive period was during the 1970s, which laid the foundation for their subsequent fame. The fourth chapter discusses Kahneman and Tversky's research of the 1970s and shows how Kahneman's psychology of cognitive mistakes provided a solution to Tversky's struggle with the *a priori* axioms of Savage's decision theory and experimental deviations from them. Kahneman and Tversky's solution was to rigorously separate the normative from the descriptive. This allowed them to maintain Savage's *a priori* axioms as the normative rules of decision making, while at the same time acknowledging the experimental results as proof that actual human decision making deviates systematically from these norms. In 1979 Kahneman and Tversky's research culminated in prospect theory, a theory which describes actual human decision behavior as a systematic deviation from the normative rules. Using prospect theory, Kahneman and Tversky deliberately broadened their scope to economics.

They considered prospect theory applicable to both economists' and psychologists' use of expected utility theory. The paper was published in *Econometrica* and argued that cognitive psychology and economic were unified in one field of behavioral science.

Chapter five investigates how economists responded to Kahneman and Tversky's understanding of experimental violations of expected utility theory and their descriptive alternative, prospect theory. It argues that there were two main responses, each with their own history. Experimental economists such as Vernon Smith corroborated and accepted the experimental results, but rejected all expected utility theories as a solution, including prospect theory. In addition, experimental economists inferred that the experimental deviations further emphasized the importance of the market as the mechanism that over time drives the economy to a rational equilibrium. Financial economists, such as Richard Thaler, also accepted the experimental results, but instead they took it as proof of the observed irrationalities in financial markets. In addition, financial economists hailed Kahneman and Tversky and prospect theory as being the most important, if not the only claimant to a solution to the problem. The use of prospect theory in financial economics led to the new field of behavioral finance. The reason for prospect theory's swift success was that it offered financial economists an elegant way out of the problems. The normative – descriptive distinction ensured that traditional, neoclassical models could be maintained as the normative theory, while at the same time it offered a descriptive alternative that was only slightly different from previously-used theories and hence easy to learn by economists.

In the late 1980s and early 1990s Thaler also started applying the behavioral finance approach to problems outside the field of financial economics. The new field grew quickly and in 1994 it was officially called behavioral economics. Once the traditional economic theories were saved in the normative realm and new theories could be developed under the rubric of descriptive theory, a surge of explorations ensued. The sixth chapter describes the history of behavioral economics in the 1990s and 2000s. Using the examples of intertemporal choice and emerging preferences it shows that behavioral economists explored many opportunities for constructing descriptive theories of economic behavior, but at the same time they always remained faithful to Kahneman and Tversky's normative – descriptive distinction. Gradually the labels of normative and descriptive were replaced by full rationality and bounded

rationality, which in turn allowed the behavioral economists to develop their own view of economic policy advice under the label of paternalism.

These developments contributed to the gradual emergence of behavioral economics as a stable and clearly defined mainstream economic program. As a result, it also brought to the fore how behavioral economists saw their program as being different from other economic programs and disciplines. Behavioral economists began to distinguish their program, in particular with regard to psychology and experimental economics. This might seem slightly schizophrenic. Behavioral economists defined behavioral economics primarily by its incorporation of psychology into economics, but at the same time they sought to distinguish behavioral economics specifically as economics, and therefore as being different from psychology. Behavioral economists relied heavily on the use of experiments and claimed the intellectual heritage of Simon, but simultaneously they explicitly distinguished behavioral economics from that other economic program that used experiments and claimed Simon: experimental economics.

How then should we understand this apparent schizophrenia in contemporary behavioral economics? In the seventh chapter I argue that the history discussed in this thesis shows how economists have actively used psychology to redefine economics. The flow of theories, methods and experimental results from psychology to economics was not a neutral process that left these theories, methods and experimental results unaffected. Instead, they lost some of their psychological connotations and gained new economic connotations. What is particularly illustrative in this regard are the two cases of experimental and behavioral economics, which both added different new economic connotations to the theories, methods and experimental results drawn from psychology. Experimental and behavioral economists used the theories, methods and results from psychology to redefine economics in their own ways. Thus, as I argue in this final chapter, this thesis not only shows that the theories, methods and experimental results that travelled from psychology to economic have not been stable entities, but it also shows that the definition of economics has not been constant. Therefore, the history of economics and psychology can only be understood by recognizing economics and psychology as disunified cultures.

2. Measuring decisions and measurement as decision in postwar psychology

1. Psychology and measurement in Michigan

From the 1950s to the 1970s the University of Michigan was the center of American psychology. It grew from seven faculty members in the late 1940s to some 225 faculty members in the second half of the 1960s [Krantz – Interview (2008), see also e.g. Peckham (2005), pp.245-266, Frantilla (1998)], and hosted the Institute of Social Research (ISR), Clyde Coombs' (1912 - 1988) Michigan Mathematical Psychology Program and Ward Edwards' (1927 - 2005) Engineering Psychology Laboratory and behavioral decision research. Over the years the ISR has received much attention in the literature [e.g. House et al. (2004), Bulmer (2001), Hyman (1991), Hollinger (1989)], while the history of Coombs' mathematical psychology and Edwards' behavioral decision research has not been fully explored. For a history of Daniel Kahneman and Amos Tversky's behavioral economics the important place to look might appear to be the ISR. Under the heading of the ISR, George Katona conducted his surveys on consumer confidence at the Survey Research Center (SRC), and even coined the label 'behavioral economics' to refer to this research. However, this chapter shows that the ISR is unimportant for the history of Kahneman and Tversky's behavioral economics, and that instead it was mathematical psychology and behavioral decision research that constituted the starting point for their subsequent collaboration.

To understand this, first the historical and organizational characteristics of the University of Michigan and its department of psychology need to be defined in more detail. Subsequently, because of its remarkable absence in the history of Kahneman and Tversky's behavioral economics, the ISR and its different centers need to be briefly discussed. After that, the third section deals with the relevant themes in mathematical psychology in the period roughly between 1950 and 1975. The fourth section describes the background and rise of behavioral decision research during the same period. Finally, the fifth section illustrates the close link between mathematical psychology, decision theory and behavioral decision research.

2. Psychology at the University of Michigan and the Institute for Social Research

David Krantz (1938-) and Robyn Dawes (1936-), two key actors in the Michigan Mathematical Psychology Program in the 1960s and 1970s, recall how the department

of psychology at the University of Michigan grew tremendously during the postwar years.⁴ In the immediate postwar years, before Coombs arrived in 1949, the department consisted of seven (voting) faculty members. As said, over the next two decades it expanded tremendously. Not all of these faculty members were full time employed by the department of psychology, although all could vote. By the late 1960s and 1970s the department of psychology employed roughly 60 full-time equivalents. Researchers had a part-time, or often even a zero-time contract with the department and held part time contracts with other institutions such as the ISR and the medical science departments. In fact, a considerable number of psychologists were working at the children's hospital, in the mental health program or in other medical science departments of the University of Michigan [Krantz – interview (2008)]. Still other psychologists were partly or wholly financed by external funds or grants. Coombs and his Mathematical Psychology Program, for instance, were financed through a grant from the National Institute of General Medical Science [Dawes – interview (2008)]. However, these multiple affiliations should not be seen as the result of vying for research funds among the psychologists. In fact, just the opposite was the case: there was enough money for nearly everyone to pursue their own ideas and interests in a general atmosphere of “live and let live” [Krantz – interview (2008), Dawes – interview (2008)]. Moreover, although the employer undoubtedly to some extent constrained the research, it was generally a non-binding way. Dawes, for instance, was employed for a year by the ISR and had an office in their building, but conducted very little work for them and continued working with Coombs and the mathematical psychologists [Dawes – interview (2008)].

These characteristics are important because it meant that if some psychologists, or groups of psychologists did not want to meet each other or discuss the merits of each other's work, they never had to because of the general availability of funds. It is in this light that the relationship between Coombs and Edwards should be seen. Both were strong, but very contrasting personalities who each had very different scientific programs, and the large number of people around and the general availability of funds ensured that they could conduct their own research programs without ever really having to confront one another. Furthermore, when two researchers with different backgrounds and research projects were interested in each

⁴ Interview of the author with Krantz, Columbia University, New York, June 20, 2008. Interview of the author with Dawes, Carnegie Mellon University, Pittsburgh, June 23, 2008.

other's work, or interested in perhaps joining forces, there was little if any pressure to do so. Thus, Coombs and the other mathematical psychologists were aware that their work concerning the axioms of measurement was in one way or another related to the measurement methods used at the ISR, and vice versa the researchers at the ISR were equally aware of the work of Coombs and others [Krantz – interview (2008)]. But in day-to-day practice both groups simply pursued their own research agendas.

In the 1950s-1970s, the department of psychology was divided into ten fields of specialization: experimental, mathematical, physiological, personality, social, community, industrial organization, and the two largest, clinical and counseling psychology. Later, physiological psychology was relabeled biological psychology and mathematical psychology became part of experimental psychology, illustrating the close connection between both. But this classification was relatively loose and more a matter of classifying what people were doing than assigning them what to do. Coombs was only associated with mathematical psychology, but Edwards' Engineering Psychology Laboratory was associated with both mathematical and experimental psychology. Tversky too was associated with both specializations. Krantz was related to experimental, mathematical, and physiological psychology and Dawes to mathematical and clinical psychology. Thus, the department of psychology had an organization, both in terms of where the money came from and in terms of fields of specialization [Krantz – interview (2008)]. But, as a result of the large number of faculty members and the availability of funds, the organization in the 1960s was not tightly knit, so that everyone could more or less do what he or she wanted to do [Krantz – interview (2008), Dawes – interview (2008)].

Related to, but organizationally distinguished from the department of psychology were the centers organized under the Institute for Social Research (ISR). The Survey Research Center (SRC) was established by psychologist George Katona in 1946, who pioneered a social survey research on consumer sentiment. To finance the war, the American government had issued a large number of war bonds and with the end of the war in sight it wanted to know how likely it was that American consumers would maintain or liquidate these bonds. Because Katona felt that he could not immediately ask people what they would do with their money, he proposed starting with some general questions that would comfort the respondents and would get him or her to start thinking about their own budgets and future prospects. In these consumer confidence surveys Katona was the first to use the term 'behavioral

economics,' as early as 1947 [Juster (2004), p.120]. Three years later, following the death of its founder Kurt Lewin (1890 - 1947) the Research Center for Group Dynamics (RCGD) was moved from the Massachusetts Institute of Technology (MIT) to Michigan. The two groups remained separate but were brought together under the newly-created Institute for Social Research (ISR). Since 1949 the ISR has been joined by other centers, and new centers have been created within the body of the ISR, such as the Center for Political Studies (CPS) and the Population Studies Center (PSC). In the 1960s, the ISR for a while contained the Center for Research and the Utilization of Scientific Knowledge (CRUSK), which later dissolved and disappeared. The scientists staffing the different centers of the ISR were social scientists and a few statisticians. Many were sociologists or political scientists, but the majority in the 1950s-1970s were the psychologists [Krantz – interview (2008)].

In order to protect its general university funds, the University of Michigan insisted upon creating of the SRC in 1946 that it was to be funded entirely through grants and contracts; a policy that was also applied to the ISR when it was created in 1949. This did not have any immediate financial implications as enough grants and contracts were available over the years. It did mean, however, that the ISR could not offer tenure to those it employed. There were always certain researchers who were the last to leave whenever funds ran out, but even these senior researchers and directors could never obtain tenure at the ISR [Krantz – interview (2008), Juster (2004), Hollinger (1989)].

The ISR and its research are remarkably *absent* in the main story of this thesis. Because of Katona's work on consumer confidence at the SCR and the term 'behavioral economics' that he created, Coombs' research and that of the mathematical psychologists is seemingly close to the psychological and social measurement of the ISR. One would imagine that there was some connection. Furthermore, Kahneman and Tversky's work during the 1960s and 1970s on human beings' perceptive and cognitive capacities, discussed in Chapters three and four, seems to be, at the very least, related to survey research on consumer confidence. In addition, one could point to the fact that Dawes was employed for a year by the ISR while working for Coombs' Mathematical Program. However, until Kahneman made a connection between his and Tversky's work and the economic and psychological survey work through his program of hedonistic psychology in the late 1990s, no link of any significance can be observed. The ISR and the research conducted at its centers

are noteworthy because of their complete absence in the history of mathematical psychology, behavioral decision research, Kahneman and Tversky's collaboration and their behavioral economics.

The reason for this is that, although both the ISR and the mathematical psychologists and behavioral decision researchers were working on psychology and measurement, in fact the two groups conducted very different projects. The ISR worked on measuring actual social, psychological, and economic characteristics of the American population. In the social psychological tradition of Louis Leon Thurstone (1887 – 1955) and Kurt Lewin it measured the attitudes of the population to spending and saving, consumer confidence concerning the performance of the economy in the near future, and so on. Mathematical psychologists and behavioral decision researchers, on the other hand, investigated the underlying characteristics of the human being regarding decision making. In their research a measurement was understood to be a human decision between two stimuli, and was thus considered to be part of experimental psychology. In a general sense both groups were working in psychology and were concerned with measurement. But their actual research was only distantly related. Kahneman and Tversky's research grew out of mathematical psychology and behavioral decision research. Therefore, the ISR is not relevant to understanding the rest of this story.

3. Mathematical psychology

The tradition of using mathematics in the study of psychological phenomena goes back to Gustav Fechner (1860) and is closely related to experimental psychology. Fechner's psychophysics was a two-sided attempt to create a mathematical basis for a scientific field of psychology and to create a mathematical basis for (scientific) measurement. As measurement occurs through human observation, a theory of human observation is at the same time a theory of measurement, and a psychological theory of observation or perception [Heidelberger (1993, 2004), Daston and Galison (2007)]. As a basis for his psychophysics, Fechner posited the idea that the just noticeable difference (jnd) is constant across individuals. For instance, the smallest increment in the brightness of a light bulb glowing at a specific brightness, at a specific distance, in a specific environment, etc., Fechner supposed to be the same across individuals. However, jnd as a basis for psychophysics eventually fell victim to its own success in the 1920s after too many jnd's had been reported and the idea of one constant jnd for

each stimulus across individuals could no longer be maintained [Gigerenzer (1987a), p.8].

Thurstone sought to save the psychophysical program in the 1930s by proposing frequency distributions instead of jnd's as a basis [Thurstone (1927a,b,c)]. Thurstone assumed that if you give two different stimuli to the individual (say two lights of different brightness) a large number of times, the relative frequency with which the individual judges the one to be larger than the other will reflect which of the two was the brighter. Moreover, and very important, when the order of objective values of the stimuli was independent of which individual perceived it, it was equally valid to ask a large number of individuals, instead of one individual a great number of times. If you wanted to know which of the two light bulbs was the brighter you could ask any individual, but if you wanted to know whether a Ferrari or a Bugatti is the more beautiful car, this method would be invalid as the order would differ across individuals. What one could ask, however, was whether drivers of a Saab consider the Ferrari or the Bugatti more beautiful, or whether Americans with a yearly income of over \$20,000 have positive or negative expectations of future economic growth, or whether Protestants consider Catholics or Muslims more benevolent. These measurements were possible when one assumed that there is one preference of *the* Saab driver for either Ferrari or Bugatti, one preference of *the* Protestant for Catholics or Muslims when it comes to benevolence, and so on.

Similar to Fechner, Thurstone's theory was as much a psychological theory of human perception as it was a theory of scientific measurement. Thurstone developed his theory of measurement to facilitate his own research on attitude measurement. In 1928, he published a small book in which he reported the results obtained from having conducted an extensive investigation on religious attitudes, investigating for instance whether *the* Protestant has an attitude to the relative importance of work and leisure that is different from *the* Catholic [Chave and Thurstone (1928)]. In a one-time attempt to extend this work to economic demand theory, after holding discussions with Chicago economic colleague and friend Henry Schultz, Thurstone sought to construct the attitudes of the individual to different combinations of hats, shoes, and overcoats. The article was published in *The Journal of Social Psychology*, but Thurstone sought to connect experimental psychology and economics by labeling the curve that connected the different combinations of goods between which the individual was indifferent an "indifference function" [Thurstone (1931)]. Thurstone

(1931) was picked up by a few economists in the 1930s-1950s [Moscati (2007)], but was, to the best of the author's knowledge, ignored by experimental, social, and mathematical psychologists.

Thurstone's measurement program was not the only existing measurement program. In the 1940s and 1950s also the representational theory of measurement rose to prominence. The most important contributor to the representational theory of measurement at this time was Stanley S. Stevens (1906 – 1973). Stevens' program was strongly inspired by Bridgman's operationalism [Bridgman (1927)], and defined measurement as the operation of assigning numerals according to a rule. Stevens distinguished between different types of measurement, ranging from the mere assignment of numerals without any further restrictions such as in the number of players on a football team, to that of ratio-measurement, in which it had to make sense to add, subtract, multiply and divide the numerals. The main question Coombs, a student of Thurstone in the 1930s, and later mathematical psychologists were interested in was whether it was possible, and if so how, to combine Thurstone's measurement approach with the representational measurement tradition.

The term 'mathematical psychology' was coined by Thurstone in the 1930s but acquired common usage in the early 1950s following the creation of Coombs' Michigan Mathematical Psychology Program in 1949. The key importance of Thurstone is always mentioned when the origins of mathematical psychology are set out [e.g. Frederiksen and Gulliksen (1964), Laming (1973), Luce, Bush, Galanter (1963a), Tversky (1991), Stevens (1951)], but the driving force behind mathematical psychology as a separate field in psychology was Coombs. An important catalyst was a two-month summer institute in Santa Monica in the summer of 1952, organized by Coombs and mathematician Robert Thrall, not incidentally a summer institute that also played an important role in shaping the newly created field of behavioral decision research and equally important in the history of game theory as revealed by historians of economics [e.g. Dimand (2005), Weintraub (1992), Lee (2004)]. The Santa Monica conference brought a range of psychologists, economists and other scientists working on the mathematical and experimental investigation of decision making together and thus facilitated the start and progress of much prominent research. Leading mathematical psychologists from the late 1950s onwards include, besides Clyde Coombs, David Krantz, and Amos Tversky, R. Duncan Luce (1925-), Patrick Suppes (1922-), and William Estes (1919-).

The contributions made to the field increased so much that in 1964 the *Journal for Mathematical Psychology* was founded.⁵ This gave self-proclaimed mathematical psychologists a more solid basis. However, it had not yet become a society. In 1975 the board of editors of the *Journal for Mathematical Psychology* discussed the possibility of a merger with the Psychometric Society and its journal *Psychometrika*. This effort was due to the financial mismanagement of *Psychometrika* and the general desire of both groups to secure their financial future by combining conferences, journal administration and so forth. But, in addition, it was argued by individual members and the board of editors of both the *Journal for Mathematical Psychology* and *Psychometrika* that also content-wise the merger might be beneficial. In the end, two proposals were put forward for voting in the two groups, one in which the two would be completely merged into one society with two journals and one in which two divisions would exist, each having their own journal under the umbrella of one overarching society. But although Coombs, Krantz, and Tversky had all indicated to Luce, one of the editors of the *Journal for Mathematical Psychology* that they would vote in favor of a merger, both proposals were rejected. In response, the editors of the *Journal for Mathematical Psychology* proposed in 1976 to create the Society for Mathematical Psychology.⁶ This proposal was accepted and the Society was officially founded in 1977.

Mathematical psychologists defined their field not on the basis of a particular understanding of psychological phenomena, but instead on the basis of a method of investigation of psychological phenomena. The field was characterized as “the attempt to use mathematical methods to investigate psychological problems,” and it was thus, “not defined in terms of content but rather in terms of an approach” [Coombs, Dawes, and Tversky (1970), p.1]. It signified “not the study of a particular type of behaviour or the delineation of some new class of psychological phenomena but, rather, the application of new techniques to traditional psychological problems” [Laming (1973), p.1]. Mathematical psychology was defined rather broadly as an attempt to use theories and techniques from the field of mathematics to represent and investigate psychological phenomena. As a result, all research that applied mathematics to what could be considered psychological phenomena in principle fell

⁵ The founding committee consisted of Richard C. Atkinson, Robert R. Bush, William Estes, R. Duncan Luce and Patrick Suppes. This paragraph draws on letters and minutes from the archive of Luce in Harvard University.

⁶ The journal editors were William Batchelder, William Estes, B.F. Green, and R. Duncan Luce.

under the heading of mathematical psychology. This is illustrated by the three-volume *Handbook of Mathematical Psychology* (1963-1965) that started its exposition of what mathematical psychology is with a list of thirty-nine “Basic References in Mathematical Psychology.”⁷ Mathematical psychology aimed to synthesize all mathematical approaches to individual human behavior.

The scope of this list of basic references turned out to be more wishful thinking than an actual reflection of research conducted by mathematical psychologists. The inclusion of economist Kenneth Arrow and political scientist Herbert Simon suggested a synthesis that did not exist. Mathematical psychology was supposed to include all mathematical reasoning related to human behavior, but in day-to-day practice it was almost exclusively focused on psychophysics, measurement theory and decision theory [Gigerenzer and Murray (1987), Coombs et al. (1970)]. Mathematical psychology of the 1950s, 1960s and 1970s was about the mathematics of measurement theory and, directly related, about the mathematics of decision theory. Decision theory will be discussed in more detail below. But before that it is necessary to devote a few words to the measurement theory of mathematical psychology.

The theory of measurement developed by the mathematical psychologists was, as said, inspired by both Thurstone’s and Stevens’ theories on measurement. Moreover, the effort to set up a mathematical psychology program by Coombs was principally influenced by Thurstone. Yet, after a while the work on measurement of mathematical psychologists drifted away from Thurstone and towards Stevens. The self-perceived task of the mathematical psychologists became to develop further the mathematical structure of Stevens’ view of measurement. The single most important publication on measurement of the mathematical psychologists were the three volumes of *Foundations of Measurement* (1971, 1989, 1990), a co-production of Krantz, Luce, Suppes and Tversky. It became the standard work on the representational theory of measurement in psychology.

In the summer of 1965, at the end of a three-week measurement workshop held at the University of Michigan, the already established scholars and long-time

⁷ These basic references include among others Arrow’s *Social Choice and Individual Values* (1951), N.R. Campbell’s *Foundations of Science* (1957), Chomsky’s *Syntactic Structures* (1957), Guilford’s *Psychometric Methods* (1954), Luce and Raiffa’s *Games and Decisions* (1957), Simon’s *Models of Man* (1957), Stevens’ *Handbook of Experimental Psychology* (1951), and Thurstone’s *Multiple Factor Analysis* (1947), and *The Measurement of Values* (1959).

friends Luce and Suppes invited the “then two brightest young people working in the area” [Luce’s letter to Hamada, June 23, 1986] to write a book on measurement that would summarize and synthesize all the recent work done on measurement in mathematical psychology. Despite the gap between the publication of the first volume and volumes two and three, most of the three volumes of *Foundations of Measurement* was written in the late 1960s.⁸

The main author of the first volume was Krantz, who consequently was also made its first author. The editor and first author of the second volume was Suppes, whereas the third volume was edited by Luce. The main initiator and contact person throughout the whole project was Luce. Luce and Tukey (1964), the very first article published in the *Journal for Mathematical Psychology*, formed the basis for much of the measurement work in mathematical psychology, and hence also formed an important basis for *Foundations of Measurement*. Interestingly, the authors discovered along the way that much of what they were doing had been done before by mathematician and economist Gérard Debreu [e.g. Debreu (1954, 1958, 1959a, 1959b, 1960)]. But Debreu had taken a topological approach that was difficult to understand for economists and psychologists [Krantz – interview (2008)]. The reference to Debreu is intriguing because it illustrates that economists and psychologists were working on the same phenomenon, but understood it differently. For mathematical economist Debreu his work was on utility theory, and for the mathematical psychologists it was about measurement.⁹

In the first two sentences of the first chapter of the first volume of *Foundations of Measurement* the authors stated their belief in the representational theory of measurement and the object of their book explicitly: “When measuring some attribute of a class of objects or events,” they argued, “we associate numbers (or other familiar mathematical entities, such as vectors) with the objects in such a way that the properties of the attributes are faithfully represented as numerical properties. In this book we investigate various systems of formal properties of attributes that lead to measurement in this sense” [Krantz et al. (1971), p.1]. *Foundations of Measurement* thus referred to the mathematical properties used in the numerical

⁸ This paragraph draws on the interview with Krantz and letters from Luce’s archive in Harvard.

⁹ Also historians of economics have focused only on the economic interpretation of Debreu’s work. For instance, Weintraub and Mirowski (1994) note that “Debreu is best read as providing a handbook for the working economic theorist of the neoclassical components of economic theory. In retrospect, it is hard to read *Theory of Value* [1959b] as anything else” (p.266).

structure in the representational theory of measurement. The first chapter puts forth what were called the three basic procedures of measurement: 1) ordinal measurement, 2) counting of units, and 3) solving inequalities. It only differs from the approach set out by Stevens (1939, 1951) in that it was more mathematically refined and sophisticated. The remainder of the book is based on these three procedures. This view of measurement served as an important component in decision theory and behavioral decision research, as set out below, but it also illustrates which approach mathematical psychologists took towards the world they investigated. I have followed the example of the measurement of length as it is used in *Foundations of Measurement* (1971). The same example was employed in Stevens (1939, 1951), and Bridgman (1927), but using less mathematical formalization.¹⁰

In ordinal measurement the only thing that is required for measuring the length of different rods is that numbers be assigned to rods of different lengths in a consistent manner. If one labels the different rods a , b etc, and considers the assignment of numbers to denote the length as a function of the rods, the only thing that is required for ordinal measurement is that “ $a \succ b$ if and only if $f(a) > f(b)$ ” [Krantz et al. (1971), p.2], in which the difference between \succ and $>$ is the difference between the empirical and the numerical structure. That is, the numerical structure $f(a) > f(b)$ can be mapped onto the empirical or natural structure $a \succ b$. A mathematical relation, here an inequality, comes to represent the relation between two natural objects, of their relative lengths in this case. Hence, if we have assigned any number to the first rod, and the second rod exceeds the length of the first rod, the only thing required in ordinal measurement is that we assign it a larger number. This is the most general and unconstrained procedure of measurement that can be applied to any attribute of any object; provided that the empirical comparison can be made and that the sensitivity of the comparison process exceeds the disparities of the objects measured.

The procedure of counting of units, which is the second procedure in *Foundations of Measurement*, is an extension of ordinal measurement that allows for a comparison to be made of the lengths of the rods. If we wish to not only represent

¹⁰ What I present here is a relatively brief sketch of one specific approach within the representational theory of measurement. For a methodological discussion of measurement in general and the representational theory of measurement in particular see Boumans (2004, forthcoming). For a thorough exposition of the history of measurement theory in nineteenth century experimental psychology and of the link of this psychological literature to interwar logical positivism see Heidelberger (1993, 2004). For a discussion of postwar measurement theory of the *Foundations of Measurement*, and its link to logical positivism/empiricism and Stevens, see Michel (1999, 2007).

that $a \succ b$, but also that, say, the length of rod a exceeds twice the length of rods of length b , hence $a \succ b \circ b'$, this is the procedure of measurement we require, where \circ is the notation for $+$ in the empirical structure and b' is employed to distinguish in the empirical structure between two rods of the same size. With respect to ordinal measurement, a number of extra assumptions are needed in order to establish this procedure for the counting of units. For example, to make the representation for the addition of $b \circ b'$ mathematically possible, we have to assume that two rods of lengths b can be represented by $2f(b)$. For the third procedure, that of solving inequalities, it requires in addition that the different distances between numbers in the numerical structure are meaningful representations for properties of the empirical structure. For instance, the numerical representation $2a + 5b = 3c$ needs to be regarded as a meaningful representation of the empirical structure.

The example illustrates that the representational theory of measurement in mathematical psychology started from mathematics and logic [Michell (2007)]. The fundamental assumption in this view of measurement is that if the scientist wants to measure, he or she needs the appropriate mathematical system. Thus, it assumed that the phenomena he or she wants to measure are clearly defined. If the scientist wants to measure length, temperature, wealth, or utility, what he or she needs to do is specify mathematically all the characteristics used in the measurement procedure *and* in the empirical system he or she wants to measure, and then afterwards apply this to the observations. When, for instance, transitivity is a mathematical requirement or a characteristic of the measurement system the scientist wants to use to measure temperature, he or she needs to start from the observation or assumption that the natural phenomenon of temperature has transitive properties. In other words, if the numerical structure that he or she uses to measure temperature has the property that it is transitive, the measurements of temperature are interpreted as transitive. This is equally true for situations where the human being is used as a measurement instrument. If the psychologist wants to measure the human perception of utilities through human beings using a measurement framework that employs transitivity, he or she needs to assume that human perception of utilities has transitive properties. Ideally one first discussed whether transitivity made sense in the case of temperature, religious attitudes or utility, but if this stage was forgotten the mathematical framework used would determine how the world was understood.

As said, from its inception measurement theory has been linked to psychophysics and experimental psychology. Heidelberger (1993, 2004) shows that from the start Fechner's psychophysics was as much a psychological theory relating objective stimulus to subjective sensation as a theory of measurement. Fechner devised his psychophysical system as a scientific foundation of measurement. It provided a scientific theory for the human body as a measurement device [Heidelberger (1993, 2004), see also Michell (2007)]. For the mathematical psychologists of the postwar period this link between psychophysics as a psychological theory and as a theory of measurement still served as the basis for their work.

It appeared to mathematical psychologists that mathematical psychology transcended the distinction between psychology and economics. As said, the basic references in mathematical psychology from the *Handbook of Mathematical Psychology* included the works of economists such as Arrow, Howard Raiffa, and Frederick Mosteller, who were considered to be important contributors to economics as well. In addition, the *Handbook* included publications written by non-economists which were considered to be important by economists for the field of economics, such as Savage and Simon. It also contained a book that was co-authored by a psychologist and an economist, Luce and Raiffa's *Games and Decisions* (1957). In addition, the summer institute in Santa Monica in 1952 provided an important impetus for both mathematical psychology and economics.

Yet, to conclude from this that in mathematical psychology economics and psychology indeed were one and the same thing, and hence unified would be a mistake. The list of basic references used in mathematical psychology contained many more books that were unfamiliar to economists than it did books that were familiar.¹¹ The 1952 Santa Monica conference, immediately mentioned when the history of mathematical psychology is touched upon, is important for mathematical psychology because it was organized by a mathematical psychologist, Coombs, and afterwards proved to have been the beginning of a rapid rise in mathematical psychological research. The mathematical psychologists did not make a link to the field of economics in relation to the Santa Monica summer institute.

¹¹ Examples include Osgood's *Method and Theory in Experimental Psychology* (1953), and Rosenbith's *Sensory Communication* (1961).

Mathematical psychology's view of economics can be further illustrated by their discussion of what economists would immediately recognize as an economic book, von Neumann and Morgenstern (1944)'s *Theory of Games and Economic Behavior*, in Coombs *et al.* (1970) *Mathematical Psychology, An Elementary Introduction*. The *Theory of Games and Economic Behavior*, Coombs *et al.* argued, is the most important modern contribution to utility theory, that is, the theory that derives from the philosophical-psychological theory of utilitarianism. It is a mathematical refinement of what is a philosophical or psychological theory. The book does, of course, have "Economic Behavior" in its title but to Coombs *et al.* economic behavior was a subset of behavior, just as social, religious, political or any other kind of behavior, and thus part of psychology. Mathematical psychologists drew on sources that economics also relied on, but they employed these sources in a different way than did economists. A similar case is Krantz, Luce, Suppes, and Tversky's reference to Debreu work as measurement theory, as mentioned above.

4. Decision theory and behavioral decision research

4.1 Decision theory

The main question that Coombs had started his mathematical psychology program with was how could Thurstone's measurement theory be brought in line with the new representational theory of measurement, a process that culminated in *Foundations of Measurement*, an axiomatic interpretation of the representational theory of measurement that has little to do anymore with Thurstone. But mathematical psychology maintained the link with Thurstone and psychophysics in general by continuing to emphasize the two-sided role of their approach as being both a theory of measurement and a psychological theory of human behavior. Furthermore, with respect to their theory of human behavior the mathematical psychologists brought their theories in line with the recent developments in theories of human behavior. The new theory they incorporated was decision theory. The thus modernized two-sided theory of psychophysics was described as follows at the beginning of Chapter eight, *Foundations of Measurement I*:

Unlike most theories of measurement, which may have both physical and behavioral interpretations, the theory of expected utility is devoted explicitly to the problem of making decisions when their consequences are uncertain. It is probably the most familiar example of a theory of measurement in the social sciences. [Krantz et al. (1971), p.369]

In mathematical psychology the originally two-sided psychophysical theory of just noticeable differences had been abandoned, but the idea of one theory serving both as a theory of measurement and as a theory of human behavior had been maintained. No economist, perhaps with the exception of Francis Edgeworth, would have understood utility theory as a theory of measurement. But for the mathematical psychologists the representational theory of measurement and the theory of expected utility theory, or decision theory, were two sides of the same psychological coin.

Decision theory studied which decision an individual should make when he or she is faced with uncertain or incomplete information. Decision theory's revival in the twentieth century was principally due to Leonard Savage. It goes back to the second half of the seventeenth century when mathematicians and other scholars started to investigate how to calculate mathematically the optimal decision in uncertain situations. The starting point is prosaically represented by the figure of the Chevalier de Méré, a notorious gentleman-gambler at the court of Louis XIV, who asked mathematicians Blaise Pascal and Pierre de Fermat to solve a number of gambling problems. The mathematics that came out of these and similar questions was probability theory and rational choice theory [Hacking (1975), Daston (1988)]. Eighteenth-century probability theory gave rise to nineteenth-century statistics and came to pervade every corner of scientific and daily life [Daston (1983,1988), Porter (1986,1994)], and it is therefore no exaggeration to characterize this development as "probabilistic revolution" [Krüger, Daston, and Heidelberger (1987), Krüger, Gigerenzer, and Morgan (1987)].

A major problem confronting probability theory was what became known as the 'St. Petersburg Paradox,' invented by Nicholas Bernoulli in 1713. Bernoulli demonstrated that gambles could be constructed for which probability theory

computed a maximum willingness to pay that was clearly at odds with intuition.¹² The most famous solution to the St. Petersburg Paradox was offered by his cousin Daniel Bernoulli in 1738 [Bernoulli (1954)]. Daniel Bernoulli distinguished between wealth and “moral wealth,” in which moral wealth depends on wealth logarithmically.¹³ Up until the early twentieth century the literature on mathematical theory of decision making under uncertainty consisted mainly of attempts to solve this and similar paradoxes [Edwards (1954), p.380].

Between the 1920s and the 1950s a number of ideas were introduced that thoroughly reshaped the way decision theorists, as they were now labeled, thought about decision making under uncertainty.¹⁴ Authors such as Bruno de Finetti [e.g. de Finetti (1937,1949, 1951)] and Frank Ramsey [Ramsey (1931)] introduced the idea that probability theory could not only be applied to objective uncertainties out there in the world, such as the probability that a coin falls heads and the probability that the sun rises tomorrow, but also to subjective probabilities, that is uncertainties inside the individual of the sort ‘how uncertain am I that it will rain tomorrow?,’ or ‘how certain am I that this second-hand car will last at least two years?’ In a related development, authors such as John Maynard Keynes [Keynes (1921)] and Rudolf Carnap [Carnap (1950)] extended the theory of logic to include uncertain propositions, that is propositions with a degree of probability that is less than 1.¹⁵ In this logical probability approach, uncertainty stems from the subject’s personal belief in the occurrence of an event. The difference between objective and subjective probability is that objective probability is a probability obtained on the basis of available information and mathematical theory, a probability that is the same for everyone. Subjective probability, on the other hand, is a number attached to the personal belief of an individual. Subjective probabilities of the same event may thus differ across individuals.

The distinction between the two was not unproblematic and this is still not the case, for it is difficult to determine where to draw the line between the two. Statistical

¹² The St. Petersburg paradox has given rise to a vast array of literature. An overview of the different sides to the debate that have developed over the past 250 years can be found in Jorland (1987).

¹³ As a synonym for moral wealth the original Latin text used the term “*emolumentum*,” which in the English translation of 1954 is translated as “utility,” upon the advice of Savage. See also Teira (2006).

¹⁴ This paragraph briefly indicates a few points in a large literature. Useful overviews are Hajek (2007), von Plato (1994), and Eriksson and Hajek (forthcoming).

¹⁵ This research can be traced back to nineteenth-century authors such as George Boole and Augustus De Morgan [e.g. Maas (2005), pp.111-122, MacHale (1985)].

data, the basis for objective probability, is information observed by human beings and can thus equally be considered input for a subjective probability. Moreover, all of the calculations for objective probability are always conducted by human beings, and can therefore also be considered as subjective probabilities instead of objective probabilities. Adherents of the so-called subjectivist or Bayesian school argued precisely this: that statistics is simply the extension of the process of human belief formation to a more formal domain. This *ipso facto* meant that the whole of statistics is a process of human decision making under uncertainty, albeit a process which is scrutinized more rigorously and recorded more formally.

In other words, the subjectivist probability theory commenced by de Finetti and Ramsey, and the logical probability approach of Keynes, Carnap and others made statistics a part of decision theory. Thus, Wald's influential *Statistical Decision Functions* (1950) stated on the first page that "[a] statistical decision problem arises when we are faced with a set of alternative decisions, one of which must be made, and the degree of preference for the various possible decisions depends on the unknown distribution $F(x)$ of X " [Wald (1950), p.2, see also Fishburn (1964)]. Decision theory was no longer only about which decision we as human beings should make given our preferences and the objective probability of different states of the world, but it was also about which conclusion should be inferred by statisticians from statistical data. Decision theory had incorporated statistics and was now an all-encompassing theory of human decision making under uncertainty.

Another new development was initiated by von Neumann and Morgenstern's *Theory of Games and Economic Behavior* (1944). In the course of constructing game theory, von Neumann and Morgenstern introduced the concept of stochastic preference, which can be found either in a weak or a strong form [see Tversky (1969) for the distinction between the two]. Stochastic preference embodies the idea that an individual who has only a very small preference for A as opposed to B, may not always correctly perceive this small difference and may mistakenly choose B. The difference is so small that he or she cannot consciously perceive it and considers him- or herself to be indifferent towards A and B. However, if the choice is repeated a large number of times, he or she will nevertheless choose A more often than B. Therefore, this individual is said to stochastically prefer A to B.

Stochastic preference eliminated the concept of indifference. Even if the individual has an infinitely small preference for A as opposed to B, this preference

would show up if the choice was repeated often enough. The individual is unaware of his or her preference for A as opposed to B and considers him- or herself to be indifferent, but he or she is not, and the mathematics therefore needs to model him or her as such. In a similar way, stochastic preference dealt the final blow to experimental psychology's just noticeable differences. Just noticeable differences, as previously mentioned, were introduced by Fechner as the lowest difference in stimulus, including that between preferences, which an individual could observe. With stochastic preference, the concept of just noticeable difference had become obsolete. The experimenter could now give the subject the same choice a large number of times and from the outcome it could be inferred which of the two options he or she preferred, even if the individual him- or herself claimed to be indifferent. Stochastic preference allowed going below just noticeable differences, and thus rendered it obsolete as a starting point for psychophysics.¹⁶

Furthermore, von Neumann and Morgenstern (1944) cut short the discussion on what exactly utility is and how it should be measured: "We [...] assume that the aim of all participants in the economic system, consumers as well as entrepreneurs, is money, or equivalently a single monetary commodity. This is supposed to be unrealistically divisible and substitutable, freely transferable and *identical*, even in the quantitative sense, with whatever "satisfaction" or "utility" is desired by each participant [von Neumann and Morgenstern (1944), p.8, emphasis added]. With regard to the unit of analysis of decision theory, von Neumann and Morgenstern thus effectively turned the clock back to before Daniel Bernoulli, when the rational decision depended on the absolute, objective value of money. For von Neumann and Morgenstern, the agents in decision problems wanted to maximize their monetary income, not their Bernoullian utility. However, von Neumann and Morgenstern (1944) labeled this money 'utility.'

The different aspects and new ideas were organized under the heading of one theory by Leonard "Jimmy" Savage (1917-1971) in his *The Foundations of Statistics* (1954). Savage divided decision theory into two realms, a normative realm and what he labeled, an "empirical" realm, a reference to the "empirical" domain of measurement theory, as discussed above. In the normative realm, rational human beings investigated how decision making under uncertainty should be done, and

¹⁶ Note the similarity with Thurstone's psychophysical theory of measurement as discussed above. See also Gigerenzer (1987a,b).

established rational principles for this behavior. In the empirical domain, scientists investigated whether people in everyday life behave according to the principles of the normative theory. For the research in the empirical domain Savage had the experimental psychologists in mind, but as a mathematician Savage himself stuck to developing the normative theory. The investigation of normative or rational decision behavior was considered a deductive science, an investigation that was best done in the comfort of the armchair. But it was, according to Savage, not just mathematicians who had contributed or could contribute to developing this normative theory. Important contributions had been made by economists and philosophers. Savage thus considered economics and philosophy to be deductive armchair sciences just like mathematics.

The purpose of *The Foundations of Statistics* was to bring together two themes in Western thought that go back to ancient Greece: inductive inference and reasoning [Savage (1954), p.1]. The formal investigation of reasoning is logic. Until the end of the first half of the twentieth century when Savage introduced his position, logic was only concerned with certain propositions; the purpose of Savage's book, and the logical probability and subjective probability tradition in which it stood, was to extend logic to uncertain propositions.¹⁷ As inductive inference typically leads to uncertain propositions, in the sense that the probability that one's inference is correct is never 1, such an extension united reasoning and inductive inference. The result was what we call statistics. "Decisions made in the face of uncertainty pervade the life of every individual and organization," Savage argued, and, "[i]t may be said to be the purpose of this book, and indeed of statistics generally, to discuss the implications of reasoning for the making of decisions" [Savage (1954), p.6].

Savage's theory investigated what a rational person does in the face of uncertainty. Rationality to Savage is a theory of reasoning, either formalized or not. For certain propositions, it is generally accepted that this theory is logic. That is, the axioms of logic are widely accepted as describing and providing rules for reasoning about certain propositions. For the extension of logic to uncertainty Savage presented in the book, this was less clear, as the theory still had to be developed. That is, Savage contended that it was as yet not clear whether the axioms he presented were indeed

¹⁷ I here forgo discussion of Savage's historical introduction. As above, one may object by pointing to for instance George Boole and Augustus De Morgan. Savage, however, does not discuss the period in between the Bernoullis and de Finetti, Andrey Kolmogoroff and Ramsey.

the best description, and provided the best rules, for reasoning under uncertainty. The reader must subsequently thus verify for him- or herself the axioms Savage presented.

How could this be achieved? As these axioms had to do with reasoning, the reader should verify them by reasoning. In fact, Savage was cautious when seeking to convince the reader of his approach. “I am about to build up a highly idealized theory of the behavior of a “rational” person with respect to decisions,” he wrote. But,”[i]n doing so I will, of course, have to ask you to agree with me that such and such maxims of behavior are “rational.” [...] So, when certain [i.e. some - FH] maxims are presented for your consideration, you must ask yourself whether you try to behave in accordance with them, or, to put it differently, how you would react if you noticed yourself violating them” [Savage (1954), p.7]. Like the axioms of logic, Savage’s “maxims of behavior” were axioms of decision making that all rational individuals should agree upon. They were independent of any preferences or beliefs and derived from an introspection that comes before experience of any kind.

Savage’s theory did not say anything about whether people in the real world actually behave according to his theory. It is important to distinguish this from the previous point. On the one hand, Savage’s readers, philosophers, mathematicians, economists, psychologists and any other rational individuals needed to investigate through introspective reasoning, whether they agreed with the new axioms for decision making under uncertainty, as they had done for over two thousand years with the axioms of logic proper. If these axioms were agreed upon, they could then be used as rules for sound reasoning, just like the rules of logic had been used as rules for sound reasoning. On the other hand, when established and agreed upon, the question could be posed whether people in their everyday decision making under uncertainty would behave in accordance with the new axioms. It should be stressed that such an exercise could only be undertaken when the rules of reasoning had been established, at the very least by the scientists conducting the empirical investigation. In other words, in such an empirical investigation into real-life decision making under uncertainty, the rules or axioms themselves were agreed to be true, and could not be experimentally scrutinized. It would nevertheless be fruitful to conduct empirical investigation in sciences, such as psychology, which were concerned with actual decision behavior by people in the real world, and not so much with the theory of reasoning itself. In order to clarify this point Savage conceptually distinguished

between the already-mentioned normative and empirical realms [Savage (1954), pp.19-20].

What conclusion needed to be drawn when a subject in an experimental setting was observed to make a decision that violated the rules governing the theory of reasoning? First of all, the theory could only be applied experimentally to subjects that can reason. Roughly, this included all normal and healthy adults; it was not useful to ask a subject that cannot reason to make a rational decision. Children, the mentally disabled and animals were therefore excluded from experimental investigation. But when subjects capable of reasoning were observed making decisions that violated the axioms, such decisions were deemed irrational decisions, or simply errors. To Savage these errors were the result of failed or too little reasoning. The individual had made a mistake in his or her reasoning or had not given it enough thought. When the subject would think further or when his or her error would be explained, he or she would recognize his or her mistake and correct his or her behavior. Savage noted that “[t]here is, of course, an important sense in which preferences, being entirely subjective, cannot be in error; but in a different, more subtle sense they can be. [...] A man buying a car for \$2,134.50 is tempted to order it with a radio installed, which will bring the total price to \$2,228.41, feeling that the difference is trifling. But, when he reflects that, if he already had the car, he certainly would not spend \$93.85 for a radio for it, he realizes that he has made an error” [Savage (1954), p.103].

How should the empirical testing of the theory be done? The central issue here was that the information needed to make a rational decision should be the same for the experimenter and experimental subject alike. The reason was that if the experimenter was not sure that the experimental subject used the exact same information as input for his or her decision, the experimenter could never establish whether the subject was making the correct decision, or an error. For instance, if the experimental subject believes that the deck of cards of the experimenter has been shuffled unfairly, while the experimenter knows that it has been shuffled fairly, the subject could make a decision that is rational given his or her own belief, but which is irrational given the experimenter’s belief. In the case of decision making under uncertainty, it should be completely clear what the uncertainty of the inference was, and what the value of the decision was. In other words, the probabilities and utilities of the different decisions involved should be clear.

To conceptually clarify, Savage invented the term ‘small world,’ as opposed to ‘grand world’ in which we live most of the time, for situations in which probabilities and utilities are clearly defined. A small world is a decision situation in which all the probabilities, utilities and consequences of the different options are clear to both experimenter and experimental subject. Therefore, “[i]t will be noticed that the small-world states are in fact events in the grand world, that indeed they constitute a partition of the grand world” [Savage (1954), p.84]. For instance, when the subject is asked to choose between five dollars for certain or a six in ten chance of winning ten dollars, this is a small world situation. The uncertainty and value of each decision are defined and clear to everyone. However, when the subject is asked to choose between a ten-year old Mercedes and a three-year old Toyota, we are in a large world decision situation. Both the value of the different options as well as the probabilities of all kinds of uncertainties associated with the two options is unclear and dissimilar for both the experimenter and experimental subject.

The value of the different options was to be measured in utilities. On the interpretation of the theory of utility, Savage fully sided with von Neumann and Morgenstern. Echoing the *Theory of Games and Economic Behavior*, Savage suggested that economists and others had been somewhat led astray in constructing complicated theories of utility, as a result of the previously-mentioned paper by Bernoulli. “For a long period,” Savage argued, “economists accepted Bernoulli’s idea of moral wealth as the measurement of a person’s well-being apart from any consideration of probability, though “utility” rather than “moral wealth” has been the popular name for this concept among English-speaking economists.” As a result, “[e]conomists were for a time enthusiastic about the principle of diminishing marginal utility, and they saw what they believed to be reflections of it in many aspects of everyday life” [Savage (1954), p.95]. However, thanks to von Neumann and Morgenstern we were now back on the right track and able to measure choice-options by using a money scale of utility. Utility equals money and is nothing more than a convenient measurement scale of preferences. “A function U that [...] arithmetizes the relation of preferences among acts will be called utility. [...] I have chosen to use the name “utility” in preference to any other, in spite of some unfortunate connotations this name has in connection with economic theory, because it was adopted by von

Neumann and Morgenstern when they revived the concept to which it refers, in a most stimulating way” [Savage (1954), p.95].

Savage’s decision theory may sound rather theoretical and anything but applicable to everyday life, but normative decision theory could also be applied to questions in the world outside science’s ivory tower. Indeed, the whole purpose of decision theory was to help us to make better decisions. During World War Two in the United States, the application of decision theory to everyday problems developed under the rubric of operations research [e.g. Klein (2000)]. Operations research aimed to gather all information available relating to a particular problem and then use decision theory to calculate the optimal decision. “Operations research makes the claim that, by pitting the forces of research against large-scale problems, the decision maker (manager, president, general, etc.) will be freed to devote his time to other tasks” [Fishburn (1964), p.4]. Although based on deductive introspective reasoning, decision theory was explicitly meant to be applied to real world decision problems. In turn, behavioral decision research was closely related to operations research. It was a newly-created field in psychology that, like operations research, sought to apply decision theory to real-world problems.

4.2 Behavioral decision research

The founding father of the empirical investigation of decision theory in psychology was Edwards, who in 1958 joined the University of Michigan [Philips and von Winterfeldt (2006)]. In Michigan Edwards founded the Engineering Psychology Laboratory to study and improve human decision making [Fryback (2005)]. Edwards was strongly influenced by von Neumann and Morgenstern’s and Stevens’ work on measurement, and was one of the first promoters of Savage’s normative-empirical decision making program. Edwards admired Savage as one admires a genius, something Edwards shared with others who had read *The Foundations of Statistics*, such as Luce, Tversky and Krantz [Krantz – interview (2008)]. Edwards’ 1954 article on the historical background of decision making research, “The Theory of Decision

Making,” and its 1961 follow-up, “Behavioral Decision Theory,” created the field of behavioral decision research.¹⁸ Behavioral decision research was dominated by Edwards until the early 1970s. From that moment on a number of his students started to develop their own interpretations. The most successful were Slovic and Lichtenstein, who developed a constructed preferences approach that drew connections with Simon [e.g. Slovic and Lichtenstein (1971, 1973, 1983)]; and Kahneman and Tversky, who created their Heuristics and Biases approach.

Edwards and his behavioral decision research adopted the framework set out by Savage, and understood Savage’s distinction between a normative and an empirical domain to be the same as experimental psychology’s distinction between normative and descriptive. Decision theory was understood as providing a theoretical framework for the objective stimuli that the subject is presented with in his case of decision making under uncertainty. The self-assigned task of behavioral decision researchers was to investigate experimentally which decision subjects make with respect to this objective stimulus. In the traditional framework, experimental psychology investigated individuals’ subjective perception of objective values, such as weight or brightness differences. In behavioral decision research the weights and light bulbs were replaced with the utilities and probabilities of decision theory. Given the objective values of the utilities and probabilities, decision theory determined the objective decision. Behavioral decision research then investigated experimentally which decision the subject actually made. In this way Savage’s decision theory and his distinction between a normative and an empirical domain were integrated into the experimental psychological framework, in which decision theory determined the objective benchmark with which the subject’s subjective decision was compared. The distinction between the normative and the descriptive was often and clearly made by behavioral decision researchers. Here is an example:

Decision theory is the study of how decisions are or ought to be made. Thus it has two faces: descriptive and normative. Descriptive decision theory attempts

¹⁸ Different names for Edwards’ program and its offspring exist. Behavioral decision research, behavioral decision theory, and behavioral decision making are all used to refer to the same psychological program. It is not clear when and how these terms exactly originated; although behavioral decision theory has been around at least since Edwards published his second overview article in 1961. Behavioral decision research, the most commonly used label seems to have originated in the 1970s, but has been applied in retrospect to the research of the 1960s also. To avoid confusion I use the term behavioral decision research in this thesis.

to describe and explain how actual choices are made. It is concerned with the study of variables that determine choice behavior in various contexts. As such, it is a proper branch of psychology. Normative decision theory is concerned with optimal rather than actual choices. Its main function is to prescribe which decision should be made, given the goals of the decision maker and the information available to him. Its results have a prescriptive nature. They assert that if an individual wishes to maximize his expected gain, for example, then he should follow a specified course of action. As such normative decision theory is a purely deductive discipline.

[Coombs, Dawes, and Tversky (1970), p.114]

It was in this regard that Edwards was interested in economics, as exemplified by his extensive and knowledgeable discussion of economics in Edwards (1954). Like Savage, Edwards understood economics as a normative, deductive theory of human decision making, and he discussed it on an equal footing with statistics, mathematics and philosophy. Thus, Edwards noted that economics is an “armchair” science [Edwards (1954), p.14], not because he denounced economics, but because he understood economics to be an armchair science just as mathematics, statistics, and philosophy. In his classification of the field of decision theory as a) the theory of riskless choice, b) the application of the theory of riskless choice to welfare economics, c) the theory of risky choices, d) transitivity in decision making, and e) the theory of games and of statistical decision function, economics is predominantly about a) and b). Von Neumann and Morgenstern’s *Theory of Games and Economic Behavior* (1944) was understood to be a deductive, armchair science as well: Game theory as a mathematical theory “can be viewed as a branch of normative decision theory” [Coombs, Dawes, and Tversky (1970), p.202].

Edwards’ discussion of “economic man” should also be read in this light. Economic man for Edwards is someone who makes his choices according to the normative theory, making it therefore a normative concept. If you ask what economic man would do in a certain decision problem, you ask what the normative solution is. At the same time, economic man as the embodiment of the normative theory, forms a hypothesis about actual decision making that can be tested: “if economic man is a model for real men, then real men should always exhibit transitivity of real choices. Transitivity is an assumption, but it is directly testable. So are the other properties of

economic man as a model for real men” [Edwards (1954), p.16]. But transitivity and the other properties of economic man were also the assumptions of measurement theory, as set out before. As a result, mathematical psychologists moved smoothly from measurement theory, to decision theory, signal detection theory and back.¹⁹ So did Edwards and his behavioral decision research.

Savage and other decision theorists investigated the normative decision theories, and it was the task of psychologists, according to Edwards, to investigate the descriptive part and in turn to see how well human beings in their actual everyday decision making behave according to the normative principles set out by decision theory. What was at least just as important for Edwards, however, was the question how human decision making could be improved. The research conducted and favored by Edwards was explicitly called “engineering psychology.” The perceived relevance of this research is illustrated by Edwards’ comments on his visit to the North American Aerospace Defense Command in the mid 1960s [Phillips and von Winterfeldt (2006), p.5]. In this command center an enormous amount of information was gathered and decisions made by the personnel had potentially enormous consequences. Therefore, it was of utmost importance not only to know how people made decisions on the basis of uncertain information, but also to find out how the decision system could be organized such that the best decision could be made. In light of future developments in the field to be made by Kahneman and Tversky, it should be noted here that Edwards and other behavioral decision theorists did not consider the human being to be an inapt or limited decision maker in the sense of not understanding the divine rules of decision theory. For Edwards, the starting point was that the human being is very capable of making complicated decisions in situations based on uncertain information. It is just that there is only so much a single human being can do. For that reason, human beings may sometimes deviate from what is normatively the right decision, and therefore it may be useful to think about how to help human beings decide when, for whatever reason, the decision making process is especially difficult or especially important.

Edwards and behavioral decision research evaluated decisions in terms of utility and extensively referred to economists and their use of the concept of utility.

¹⁹ Signal detection theory (SDT) is a branch of psychophysics that investigates the individual’s ability to distinguish between signal and noise. In other words, it investigates decision making under noisy conditions. See e.g. Green and Swets (1964).

Nevertheless, Edwards and behavioral decision research did not understand utility in the same way as economists. For behavioral decision research utility was merely a new concept for an already existing idea in experimental psychology, that of valence. “The notion of utility is very similar to the Lewinian notion of valence. Lewin conceives of valence as the attractiveness of an object or activity to a person. Thus, psychologists might consider the experimental study of utilities to be the experimental study of valence, and therefore an attempt at quantifying parts of the Lewinian theoretical schema” [Edwards (1954), p.25, see also Frijda (1986)]. Valence measures the intrinsic attractiveness or averseness of an individual to a certain event, object or situation. Thus, if an individual is more attracted to Islam than to Christianity, Islam has a higher valence. In addition, emotions can be classified in terms of valence. Anger and fear are emotions with a negative valence, joy has a positive valence. By equating utility with valence, Edwards and behavioral decision research understood utility to be a general measurement of an individual’s attitude towards events, objects and situations. As a result, an individual preferring ten to eight dollars, was psychologically in the same situation as an individual preferring Islam to Christianity.

In behavioral decision research, the behavior of the experimental subjects was evaluated in terms of the normative benchmarks. The human being was considered to be a mechanism that reasons logically and applies Bayesian statistics. In other words, the individual was considered to be a logician and Bayesian statistician of some sort. The purpose of behavioral decision research, then, was to figure out whether this human being is a good logician and Bayesian statistician. This particular type of understanding of human behavior was neatly summarized in a paper by Rapoport and Tversky. “[The behavioral decision research] approach to the study of choice behavior,” they argued, “is based on the comparison between the normative solution of a decision problem and the observed solution employed by subjects.” As a consequence, “man is viewed as an intuitive statistician who acts in order to maximize some specified criteria while operating on the basis of probabilistic information” [Rapoport and Tversky (1970), p.118].

Edwards and the developing behavioral decision research approach created a program that took decision theory as provided by mathematicians, economists and philosophers, and especially Savage, as point of departure. It compared actual human decision making with respect to this norm, measuring the decisions made in terms of “utility,” and looked for ways to improve human decision making. But behavioral

decision researchers recognized that matters were a little more complicated than they usually portrayed them. In their introduction to *Decision Making* (1967), for instance, Edwards and Tversky noted that “the distinction between what an organism should do and what it does do is slippery” (p.8). The problematic distinction between the normative and the descriptive was a recurring theme, although it was far outnumbered by the instances in which the distinction was standardly used. The problem was that in Savage’s decision theory the normative and the descriptive were closely related. The normative rules were rules that every healthy adult should agree with when thinking them carefully through. The normative decision theory was as much a prescription for optimal behavior as it was a description of an adult’s behavior who has carefully thought through which decision to make. In experimental psychology, however, the distinction was much stronger. In experimental psychology, the descriptive value of the stimulus, the sensation, was supposed to deviate from the objective norm. Thus, when decision theory was integrated into the experimental framework, the normative-descriptive distinction of decision theory risked becoming a much stronger and much more absolute distinction than it was meant to be. This was unproblematic as long as the experiments showed that most of the time subjects indeed did make their decisions according to the norms of decision theory, and it was what Edwards and his behavioral decision researchers expected to find and actually did find. However, when the experiments indicated that there might be systematic differences between the norms of decision theory and actually observed behavior, an idea that gradually developed during the 1960s (treated in the third chapter), it did become problematic.

Throughout his career Edwards wanted to maintain the initial decision theoretical understanding of the close connection between the normative and the descriptive. Until the early 1970s, his disciples in behavioral decision research kept this perspective as well. Normative theory described human behavior in situations where we really want to behave as best as we can, for instance in cases where the stakes are high. Normative theory was thus to some extent descriptive. Moreover, “[d]ecision theory may be viewed as primarily an analysis of the environment; that is, an orderly summary of those features of the environment that control behaviour.” Therefore, “[s]uch a description of the environment, combined with the simple assumptions about behaviour tendencies that the organism brings to that environment, may yield an effective description of behaviour” [Edwards and Tversky (1967), p.8]. Although the distinction between normative and descriptive was used all the time, it

was at the same time clear that the two sides were closely connected and that perhaps it was not possible even to distinguish between the two.

Decision theory and behavioral decision research in the 1950s and 1960s both considered themselves to be directly related to economics and used extensive amounts of economics. It is especially Edwards' evidently extensive knowledge of economics [e.g. Edwards (1954, 1961)] which tempts the reader to conclude that here we have a case in which psychology and economics were truly integrated into one research project. But the way in which Savage, Edwards and others talked about economics does not resonate with the way in which economists spoke about economics. Such different prominent economists as Lionel Robbins, Paul Samuelson, and Milton Friedman would not have agreed to be engaged in constructing a normative theory of decision making.

Some of this incommensurability showed up in psychologists' assessment of economics. In his discussion of Samuelson's economics, Edwards was somewhat puzzled that "[i]f preference is operationally defined as choice, then it seems unthinkable that this requirement can ever be empirically violated" [Edwards (1954), p.15]. Moreover, the interpretation of utility in terms of Lewinian valence appears, if perhaps not entirely incompatible, not exactly what economists had in mind when they use the concept of utility. Thus, although the frequent references to economics in decision theory and behavioral decision research suggest otherwise, economists and psychologists understood their disciplines and the relationship between them in fundamentally different ways. Quite a few theories and concepts traveled from economics to psychology. But the way in which these theories and concepts were used in psychology was not something economists would have recognized as belonging to their field.

5. "Measurement theory in psychology is behavior theory"

Mathematical psychology was directly related to decision theory and behavioral decision research. Mathematical psychology applied mathematics to the investigation of psychological phenomena, and as both decision theory and behavioral decision research used a great deal of mathematics, a natural and direct link existed between the two. How to formulate mathematically how people should behave and how people actually do behave in situations under uncertainty, were research questions that belonged to mathematical psychology as well as to decision theory and behavioral

decision research. Hence, the same scientist could naturally be perceived as being a contributor to these different fields at the same time. Tversky, Luce, and Suppes serve as examples.

But the link between mathematical psychology, decision theory, and behavioral decision research also went much further than the mere use of mathematics. Mathematical psychology's representational theory of measurement and behavioral decision research's experimental investigation of human decision making started from different perspectives, but were partly about the same subject: normative decision behavior. Mathematical psychology's representational theory of measurement used the human body as a measurement device. In the case of utilities and probabilities, for instance, the human being was used to measure human perception of utilities and probabilities, human perception of risk averseness, and human perception of loss averseness. But in order to make this a valid procedure it must be assumed that the human being as a measurement device functions consistently. Furthermore, the representational theory of measurement's definition of consistency was: according to the normative rules of decision theory. The assumption needed to be made was that the human measurement instrument behaved according to the normative decision theory.

Behavioral decision research, on the other hand, compared behavior of individuals in its experiments with the norms of decision theory, for which it used the representational theory of measurement. The two fitted neatly together. Assuming that subjects behave according to the normative rules, the mathematical psychologists set up measurement frameworks that measured the perception of utilities, risk averseness and so on. Assuming that, in general, subjects behave according to the normative rules, behavioral decision researchers investigated under which circumstances subjects made mistakes. Mathematical psychologists provided behavioral decision research with a solid theory of measurement, and behavioral decision research informed mathematical psychologists under which circumstances its human measurement instrument was less accurate.

To illustrate further why for mathematical psychologists "measurement theory in psychology is behavior theory," [Coombs (1983), p.36] it is useful to ask how experimental psychologists measured the phenomena they were interested in. How did they measure the attitude of religious people who go to church twice a day? How did they measure the perception of "rape" in terms of good versus bad? How did they

measure the perception of a probability of 0.01%? How did they measure the relative utility of receiving a certain ten dollars as opposed to a 0.8 chance of receiving fifteen dollars? The answer, as already indicated, is that they measured all these psychological phenomena through the human being. “In psychological measurement, the individual is the measuring device; he plays the role of the pan balance, the meter stick, or the thermometer” [Coombs (1983), p.36]. The psychologist used individuals to measure the value of psychological phenomena of *the* individual. This could be the human being in general, it could be the member of a culture, and it could even be the individual itself. One could, for instance, use individuals as a measurement instrument to measure the individual’s risk averseness. “Psychological measurement theory is concerned with the empirical regularities in [the individual’s] behavior that justify numerical assignments to the stimuli he is responding to and/or justify numerical assignments to him” [Coombs (1983), p.36].

However, to “justify numerical assignments” to stimuli and to “justify numerical assignments” to the individual on the basis of “empirical regularities in this behavior” the psychologists needed to understand that behavior. They, in other words, needed a theory describing human behavior. The psychologists needed to understand how humans function to be able to use them as measurement instruments, just as the physicist needs to understand how the thermometer works in order to use it as an instrument. But in the case of the human being as a measurement instrument, this could not be just any understanding; it needed to be a rational understanding. Work done by Heidelberg (1993, 2004) points us to the fact that in nineteenth-century German experimental psychology, the human being functioned as a measurement device. We now see that post World War Two work regarding the representational theory of measurement and in behavioral decision research showed that in order for the human being to function as a measurement instrument the human being needed to be understood as behaving rationally. The psychologist needed to have a psychophysical or decision theoretical explanation of the individual’s response towards different stimuli in terms of *rationality*. In the case of decision making on the basis of utilities and probabilities, that theory of rationality was decision theory. Decision theory explained how an individual would rationally respond to different stimuli, and thus informed the psychologist which numeral to assign to the different stimuli. To make the link between measurement theory and decision theory one had to

assume that the individual that is used as the measurement instrument behaves rationally.

What happens if we find out that this individual in fact does not always behave rationally? This question did not come up seriously until the late 1960s and will be dealt with extensively in later chapters. However, from the above we can see what happens. If individuals are found to behave irrationally, this is problematic for decision theory because it means that decision theory does not provide a good description of human behavior. As long as the deviations from decision theory are random this is not too problematic. It would be the same problem as knowing that some or even all of the thermometers do not measure exactly right but that they measure correctly on average. However, when individuals are found to deviate systematically from the norms of decision theory, it becomes a serious problem. It not only means that decision theory is not a good description of actual, rational human behavior, it also implies that measurement theory is based on flawed assumptions. For instance, if the psychologist wants to measure what the relative value of two uncertain outcomes is and assumes that people have decided rationally, he or she simply asks a few people which of the two they prefer and thus measures which of the two has the highest expected value. But if it now turns out that human beings systematically deviate from rational behavior, the psychologist cannot infer from their choices, i.e. from the measurement, which of the two options has the higher expected value.

6. Conclusion

Mathematical psychology continued experimental psychology's focus on mathematization and measurement. In the postwar period it aligned itself with the re-appearance of decision theory in the work of Savage, and with the empirical investigation of decision theory in Edwards' behavioral decision research. This alliance proved that in order to use the human being as a measurement instrument in psychology, it needed to be assumed that the individual makes its decisions rationally. Behavioral decision research was related to mathematical psychology's measurement theory in its use of measurement theory. Behavioral decision research compared experimentally actual human decision behavior with the norms of decision theory, with the explicit purpose of engineering solutions for situations in which decision making is particularly difficult, or the individual is prone to make mistakes. The three intertwined developments of mathematical psychology, decision theory, and

behavioral decision research together constituted a scientific program of human decision behavior revolving around a set of axioms that determine rational or normative decision behavior. Furthermore, a comparison of human behavior to this normative benchmark could be made within a descriptive domain by means of experimental investigation.

It is tempting to conclude from the many references made to economics in mathematical psychology, decision theory and behavioral decision research that the three were connected to economics. And to some extent this is true. Mathematical psychology did incorporate economic texts, and Edwards, and to a minor extent Savage, based their research on extensive discussions of economics. But mathematical psychology, decision theory, and behavioral decision research used the economic literature for their own purposes, and they did this in ways that were at odds with economic practice.

3. Tversky's behavioral deviations and Kahneman's cognitive mistakes

1. Kahneman and Tversky before Kahneman and Tversky

The previous chapter showed how in postwar American psychology measurement theory was central to the development of mathematical psychology. This was further reflected in the newly-created field of behavioral decision research in psychology that assigned itself the task of testing decision theory's axioms of normative behavior experimentally. This was done with the purpose of engineering solutions for cases in which the human decision making machinery failed. Because mathematical psychology's measurement theory used the human being as a measurement instrument, and because behavioral decision research required a measurement instrument to conduct its experiments, mathematical psychology and behavioral decision research were natural extensions of one another. What was important for their interrelationship was furthermore that both used the University of Michigan as their principal base, although contributors to both programs also came from other universities.

In 1965, Amos Tversky (1937-1996) obtained his PhD from the University of Michigan under the supervision of Clyde Coombs and Ward Edwards. Tversky embodied the close connection between mathematical psychology and behavioral decision research. His dissertation combined Edwards' focus on normative rules for rational behavior with Coombs' interest in descriptive theories of measurement. Within a few years after finishing his PhD, Tversky became one of the foremost contributors to mathematical psychology's representational theory of measurement and was one of the four authors of *Foundations of Measurement*, the standard work in the representational theory of measurement. Besides his work in mathematical psychology Tversky collaborated with Edwards and quickly became a frontrunner in behavioral decision research. Tversky's prominence in behavioral decision research is exemplified by the fact that the collection of core behavioral decision research publications, *Decision Making* (1967), was co-edited by Edwards and Tversky.

Tversky's greatest claim to fame came in the 1970s from his work with Daniel Kahneman (1934-). Together Kahneman and Tversky constructed a new approach to human decision behavior in behavioral decision research that soon became more prominent than the approach favored by Edwards. They termed the new approach 'Heuristics and Biases.' Heuristics and Biases and its derivative prospect theory laid

the basis for behavioral economics which will be dealt with in Chapter four. Before we look at Kahneman and Tversky's collaborative research in the 1970s, however, we need to understand why Tversky became dissatisfied with the framework of behavioral decision research in the 1960s. To understand Kahneman and Tversky's approach to behavioral decision research of the 1970s, it is necessary to understand the problems that Tversky tried to solve. The first section of this chapter thus deals with Tversky's work in mathematical psychology and behavioral decision research in the 1960s.

But Kahneman and Tversky's work of the 1970s was not solely the product of Tversky's mind. Indeed, it is my belief that the crucial twist that solved Tversky's problems in behavioral decision research came from Kahneman. An experimental psychologist just as Tversky, Kahneman came from a different theoretical background. The research that Kahneman conducted during the 1960s included the psychophysics of vision, semantic differentials, trade-offs between different cognitive tasks and related issues. The recurring theme in Kahneman's work in the 1960s, and in his whole career, is that of cognitive mistakes. In all his research Kahneman investigates when human beings make cognitive mistakes, how severe these mistakes are, what explanation might account for this mistaken decision making, and in what way can mistakes be prevented in the future. In Kahneman's work experimental psychology's emphasis on deviations from the norm behavior became strongly pronounced. To understand why and in what way Kahneman altered Tversky's behavioral decision research, we therefore need to go back to Kahneman's work in the 1960s. The second section of this chapter deals with Kahneman's work in the 1960s.

2. Caught between *a priori* axioms and behavioral deviations: Tversky's research in the 1960s

Between finishing his PhD in 1965 and completing his first publication with Kahneman in 1971 Tversky worked simultaneously in three related areas. He worked within Leonard Savage's decision theory on the further refinement of normative decision theory, he extensively contributed to the mathematical development of the representational theory of measurement in mathematical psychology, and he measured the actual decision behavior of human beings in experiments and compared these measurements to the norms of decision theory. In the section which follows, I will

discuss Tversky's work in these three areas separately, before making an overall assessment of Tversky's work during the first six years following his PhD dissertation.

2.1 *Decision theory*

In his work on decision theory in the 1960s, Tversky adhered to the model and approach set out by Savage: "Utility theory, or the subjective expected utility (SEU) model, is a theory about decision making under risk. It is based on a set of seemingly reasonable axioms (Savage, 1954) which imply that an individual's choices between risky alternatives can be described as the maximization of his subjective expected utility" [Tversky (1967c), p.27]. Tversky also accepted Edwards' interpretation and application of Savage's decision theory in terms of psychologists' normative-descriptive distinction. Thus, we read that utility theory "has been widely applied as a normative principle in economics and operations research, it underlies game theory and detection theory, and it has stimulated extensive experimental investigation" [Tversky (1967c), p.27].

For instance, in Tversky's first article "On the Optimal Number of Alternatives at a Choice Point" (1964), he described a mathematical model that determines the optimal number of alternatives at a choice point in a test. It considered as examples choices in "[m]ultiple choice tests, mazes or personality checks" [Tversky (1964), p.386], but the argument made implied that the theory could be applied to any choice problem in tests. The paper's model showed mathematically that the optimal number of alternatives at a choice point in terms of discriminability, power and information is three. Tversky applied the approach of decision theory to a specific problem that as yet had not been solved satisfactorily, and he subsequently constructed a norm for this specific problem. Applying decision theory to the number of options at a choice point in a test yielded an optimum of three. A rational professor should thus give his students multiple choice tests with always three alternatives.

For many decision situations it had already been established how a rational individual should behave. But often the mathematics of the solution to these decision situations could be improved. The theory was there, but it needed to be worked out a bit more. In the brief theoretical part of Tversky and Russo (1969), for example, the authors investigated the fundamental principle underlying "probabilistic theories of choice behavior" [Tversky and Russo (1969), p.1]. This fundamental principle, the

authors argued, appeared in different parts of the literature in different forms. It was known as the assumption of simple scalability, strong stochastic transitivity, substitutability, or independence. The authors showed that when these four principles were formulated in mathematical terms they amounted to the same. Tversky and Russo showed how what appeared on the surface to be different types of psychological investigations, in fact share a similar mathematical structure.

Tversky's work on decision theory in the 1960s accepted the decision theoretical work of Savage in particular as a starting point and refined it by applying it to specific situations and linking it to other theories. In his work on decision theory, Tversky was an exponent of the mathematical psychology in which he obtained his PhD. Tversky's purpose was not to come up with a new theory or to criticize a theory. Instead, the focus was on refining mathematically what was already there.

2.2 Measurement theory

A major part of Tversky's research during the period 1964-1971 was on measurement theory. Five lengthy articles in this period discuss topics such as the development of a generalized model of conjoint measurement, which allows one to measure probabilities and utilities in one and the same experiment; the foundations of multidimensional scaling; and multidimensional representation, which looks for ways to decompose different dimensions of measured dissimilarity judgments.²⁰ The main question of this research, as set out in Chapter two, was to investigate which mathematical structure the measurement procedure requires so as to measure what it should measure. An important issue was how the experiment should be conducted so that both the subjective probability and the subjective value could be derived from one and the same choice of the subjects in the experiments.

In this measurement literature Tversky equated measurement theory with decision theory. Tversky was a prominent contributor to measurement theory in mathematical psychology and his views, in this regard, were the same as the mathematical psychology community at large, as depicted in Chapter two. To Tversky, the individual in psychological experiments served as the measurement

²⁰ In the preface of *Foundations of Measurement* (1971), Krantz, Luce, Suppes, and Tversky acknowledge that all the articles on measurement that appeared in the period 1967-1970 by one, or a combination of the authors, were by-products of work on the book. That is, what is in these articles can also be found in the book. The five articles on measurement written by Tversky before 1971, four of which are co-publications with Krantz, were all written during these four years.

instrument, just like the thermometer or pan balance in the experiments of the physicist and chemist. As a consequence, the axioms of the representational theory of measurement that described the working of the measurement instrument were exactly the same as the axioms that described rational, normative behavior in decision theory. As this is a crucial part of Tversky's psychology, I will quote again from *Foundations of Measurement*.

Unlike most theories of measurement, which may have both physical and behavioral interpretations, the theory of expected utility is devoted explicitly to the problem of making decisions when their consequences are uncertain. It is probably the most familiar example of a theory of measurement in the social sciences. [Krantz et al. (1971), p.369]

To Tversky, the axioms of normative decision theory were at the same time axioms that described the functioning of the psychologist's measurement instrument and axioms that described optimal decision behavior. Conceptually, this is only possible when one assumes that actual human decision behavior deviates very little from the norms of decision behavior. It assumes that if actual decision behavior deviates from the norm, it is somehow distributed evenly around the norm and does not deviate from the norm systematically. It also assumes that if human beings are found to deviate from the norms in certain situations, their mistakes can be relatively easily explained to them after which the individual will correct his or her behavior. Thus, when Tversky developed a conjoint measurement model for the representational theory of measurement using one measurement instrument to measure both utilities and probabilities, he needed to know how the subjects/instrument would behave in the experiment. For Tversky, the axioms of the measurement theory and of decision theory would predict how the subject/instrument would behave. On the basis of this knowledge, Tversky could then devise a measurement model and experiment that would produce both the perceived utilities and the perceived probabilities.

2.3 Behavioral decision research

Tversky's work on the representational theory of measurement and on decision theory came together in his experimental work on behavioral decision research. The measurement models were applied in experimental testing and actual behavior was

tested against normative decision theory. It is important to emphasize that in this early experimental work of Tversky, the question was *not* whether, let alone how, human beings deviate from the norms of decision theory. Tversky's basic research question in the 1960s, just as that posed by Edwards, was how to apply the normative model to human decision making behavior. The axioms of decision theory indicated how we should, and usually do make decisions given the utilities and probabilities implied by the different options. But what the axioms neglected to specify was how a human being perceives utilities and probabilities. In Savage's subjective expected utility (SEU) model, for instance, it was assumed that subjects have a subjective perception of both value and probabilities, termed utility and subjective probability respectively, which differs from the objective values and that subjects base their decisions on these subjective values. In a closely related model, the subjective expected value (SEV) model, it was assumed that subjects have only a subjective perception of the probabilities. Furthermore, it was investigated whether the subjective value curve of different goods has the same shape. It could, for instance, be the case that the subjective perception of the utility of candy decreases much faster than that of cigarettes. Thus, Tversky's experiments were first and foremost an investigation of how to apply the axioms of measurement theory and decision theory. The human being was used as a measurement instrument to measure the different attributes of actual human decision behavior, which in turn informed Tversky how decision theory best fitted into the experimental psychological framework.

At the same time, Tversky used his experiments to test whether the human instrument indeed functions properly. Tversky's experiments were set up to measure the subjective value curve of candy and cigarettes, but the experiments at the same time checked whether the subjects behaved according to the axioms of decision theory and measurement theory. As, for instance, the axiom of transitivity was the "cornerstone of normative and descriptive theories," and underlied "measurement models of sensation and value" [Tversky (1969), p.31], the experiments were used as an opportunity to also test the axioms of measurement theory. Thus, in one and the same experiment Tversky would apply the representational theory of measurement, check whether its measurement instrument functioned properly, measure subjective perception of probabilities and utilities, and monitor whether human beings indeed behave according to the axioms of decision theory.

For example, Tversky (1967a) tested the additivity and the independence axiom, two key axioms of normative decision theory and measurement theory, in a gambling experiment with eleven male inmates at the State Prison of Southern Michigan. The subjects had to gamble for, and were paid in, candy and cigarettes. Six normative decision models were compared for both the set of candy gambles and the set of cigarettes gambles. In both cases, it turned out, Savage's normative SEU model provided the best description of the behavior displayed by the subjects. Given the SEU model, both additivity and the independence of subjective probability and utility were confirmed. That is, assuming that people make their decisions according to the normative theory, the SEU model provided the best description. Nevertheless, Tversky was cautious and concluded that "After more than 15 years of experimental investigation of decisions under risk, the evidence on the descriptive validity of the SEU model is still inconclusive" [Tversky (1967a), p.199].

In a follow-up paper, Tversky (1967b) set out a measurement model that tested the descriptive validity of different normative models of decision making, among them Savage's SEU model, the power utility theory, and the strict additive model. To do so, the eleven inmates from the State Prison of Southern Michigan had to choose between different gambles, but did not know the relevant probabilities beforehand. Savage's SEU model provided the best description, but failed in the sense that the subjects consistently overestimated low and underestimated high probabilities. The model was therefore extended with a power utility function that allowed utility to be a non-monotonic function of money (versus monotonic in the standard case), and to vary across individuals. Tversky stressed the proven independence of subjective probability and utility in the experiment and concluded that the best descriptive model (SEU plus power utility) was incompatible with utility theory. In other words, to maintain the assumption that subjects make their decisions in the normatively correct way, Tversky had to assume a normative model that is inconsistent with utility theory. To maintain the idea that individuals make decisions rationally, it had to be concluded that utility theory is descriptively wrong.

Thus Tversky went a step further than Edwards in testing the axioms of measurement theory and decision theory. For Edwards the axioms were *a priori* truths that could not be violated. If the experimental results indicated that the axioms had

been systematically violated, then there must have been a problem with the experimental design somewhere. During the period between 1965 and 1971, however, Tversky came to believe that the axioms were systematically violated by the experimental subjects. In his experimental work, Tversky consistently distinguished between the normative and the descriptive. The emphasis was on well-known decision problems such as making gambles, but other decision situations that were investigated were, for instance, whether people can determine which of two lights is the brighter, or from which of two distributions a four-digit number can be drawn. Without exception, the experiments tested one or more normative models descriptively. That is, the hypothesis was that the normative model was a good description of actual decision behavior, which was then tested experimentally.

In a collaborative paper Tversky and Edwards (1966) investigated whether subjects seek the optimal amount of information that normative decision theory predicts that they will seek. To test this, an experiment was conducted in which subjects had to determine which of two lights is brighter. Subjects could obtain information about their results by paying with the money they had earned by giving correct responses. In some of the experimental treatments the subjects were told beforehand the distribution of each of the two lights. The normative model that was proposed as a descriptive model did not work well; subjects sought much more information than the model had predicted they would. The normative model which served as description of actual choice behavior was subsequently rejected. But the authors did not really know what to conclude from these results. They did not draw the conclusion that the normative model had been falsified, but they provided a number of explanations that might account for the deviations. At the same time, however, it was stressed that these explanations only partially explained the deviations. They concluded that for reasons yet to be discovered the normative model did not work well in this particular situation.

The problem of systematic deviations continued to bother Tversky and in 1969 he published a paper, entitled “Intransitivity of Preferences,” in which experiments were described and discussed that had the sole purpose of testing the axiom of transitivity. Transitivity, Tversky stated, “is of central importance to both psychology and economics. It is the cornerstone of normative and descriptive decision theories.”

Furthermore, it is the essential assumption in measurement theories since “it is a necessary condition for the existence of an ordinal (utility) scale” [Tversky (1969), p.31]. The article described a number of experiments that falsify weak stochastic transitivity (WST). In WST, transitivity of preferences is defined in terms of probabilities, hence, x is weakly preferred over y if and only if $P(x,y) \geq \frac{1}{2}$, meaning that the probability of choosing x over y is larger than or equal to a half. It was shown that WST does not hold descriptively. That is to say, the subjects’ actual decisions were not in the least stochastically transitive. And as transitivity is a key assumption in decision theory and measurement theory, this was potentially a serious problem as it implied that no normative model whatsoever could describe subjects’ actual decision making.

But Tversky was still reluctant to give up on this foundation of both measurement theory and decision theory, as transitivity “is one of the basic and the most compelling principles of rational behavior” [Tversky (1969), p.45]. Tversky suggested that normative decision theory could be maintained because apparent intransitivities could always be attributed to an unobserved change of preference that takes place between the decisions made. He concluded somewhat paradoxically that “The main interest in the present results lies not so much in the fact that transitivity can be violated but rather in what these violations reveal about the choice mechanism and the approximation method that govern preference between multidimensional alternatives” [Tversky (1969), p.46].

The reason that Tversky was reluctant to accept that the experiments falsified the axioms was that the axioms were such a fundamental aspect of both measurement theory and decision theory. If he had accepted that the axioms were wrong, the representational theory of measurement and normative decision theory as description of human behavior would be falsified. Another reason why Tversky was reluctant to give up the axioms was that the experimental results did not always indicate falsification. In the above-mentioned experiments conducted in the prison of Southern Michigan, the decision behavior of the subjects largely corresponded to the norm. There were more situations in which the results were mixed. In Rapoport and Tversky (1970), subjects were presented with a sequence of offers and they had to decide at each stage whether to stop and take the present offer, or to continue sampling more

offers. The optimal stopping point was given by the normative analysis. It can be shown mathematically that with, for example, 200 offers and zero costs it is optimal to let 74 offers pass and then pick the first offer that is higher than any of the offers encountered before. The authors found that in about one-third of the cases the subjects made decision ‘errors’, i.e. deviations from the norm. But in roughly two-third of the cases the subjects’ behavior was in agreement with the normative model.

Over time it became clearer to Tversky that normative decision theory was often too difficult to apply to actual human decision behavior. Although the normative theory was not always violated and could often be saved as descriptive theory by means of *ad hoc* assumptions, the pressing conclusion was that in too many cases people’s actual decisions systematically deviate from the optimal decision as determined by measurement and decision theory. Decision theory could only be proven incorrect by *a priori* introspective reasoning (as set out in Chapter two) but persistent deviating decision behavior by people in experiments could nevertheless be an indication that something was wrong with the theory. If one observes the reasoning behavior of subjects which deviates from what is considered to be logically correct, either these subjects do not reason logically, or the assumed theory of logical reasoning is incorrect.

2.4 Situating Tversky’s experimental method

Tversky’s increasing conviction that often the normative theory could not be applied to actual human decision behavior was the result of the experiments he conducted in the 1960s. These experiments were done with a small number of subjects; seven or eight was a normal group size. Although this was not always explicitly indicated, it is furthermore clear that in the majority of cases the experimenter and subject were previously acquainted since, for example, the subjects had participated in a university course the experimenter taught as their professor.

The experiments were done in a traditional experimental psychological setting. The subjects were often assigned numbers, for instance 1-7, and referred to individually. When discussing the empirical results, the subjects were sometimes analyzed individually, which was typically exemplary for a perceived more general

behavioral pattern. Thus, for instance, in an experiment published in 1969, the utility curve of subject 3 was discussed because of its peculiar shape, and in similar experiments conducted in 1966, statements of a post-experiment interview with subjects were compared with their performance during the experiment [Tversky and Edwards (1966), Tversky (1969)]. In this regard, Tversky's experimental work is an example of research from before the "inference revolution" of the mid-twentieth century [Gigerenzer and Murray (1987), p.182, see also Danziger (1990)]. That is to say, the analyses did not calculate an average response over experimental subjects, but instead tried to find an explanation that would cover the observed behavior of the individual experimental subjects.

The individual trials of the experiments were relatively long. For example, in one experiment [Rapoport and Tversky (1970)] subjects were asked to judge which of two light bulbs was brighter one thousand times in a row. The experiment consisted of a sequence of two or three trials of about one hour each. As a result, the experiment took quite some time. In this most time-intensive experiment done by Tversky, "[t]he subjects met five times a week for seven weeks. Each experimental session lasted about two hours" [Rapoport and Tversky (1970), p.108].

2.5 The road not taken: Elimination by Aspects

One conclusion Tversky was aiming to develop from 1966-1967 onwards was that people do not behave according to the normative theory of decision making, but nevertheless act rationally [Krantz – interview (2008)]. In other words, Tversky tried to develop a new normative theory. "Elimination by Aspects: A Theory of Choice" appeared as a monograph in 1971 and as an article in *Psychological Review* in 1972. "Elimination by Aspects: A Theory of Choice" began by introducing decision theory's two-fold problem with the independence axiom. The independence axiom was first of all problematic on empirical grounds because it was "incompatible with some observed patterns of preferences which exhibit systematic dependencies among alternatives" [Tversky (1972), p.281]. Even though he had observed this problem before, Tversky now concluded that this behavioral deviation could not be solved within existing decision theory:

data show that the principle of independence from irrelevant alternatives is violated in a manner that cannot be readily accounted for by grouping choice alternatives. More specifically, it appears that the addition of an alternative to an offered set “hurts” alternatives that are similar to the added alternatives more than those that are dissimilar to it” [Tversky (1972), p.283].

Because of the impossibility of solving the behavioral deviations within the existing theory, decision theorists and behavioral decision researchers required “a more drastic revision of the principles underlying [the] models of choice” [Tversky (1972), p.283].

In the theory as constructed by Savage (1954) only deductive, introspective reasoning could show the normative theory to be wrong. Empirical results could not show these rules of logical, rational reasoning to be false. Tversky accepted this position but moved on to show that on the basis of deductive armchair thinking doubts could also be raised. “Suppose,” Tversky argued, “you are offered a choice among the following three records: a suite by Debussy, denoted D, and two different recordings of the same Beethoven symphony, denoted B_1 and B_2 .” Assume furthermore “that the two Beethoven recordings are of equal quality, and that you are undecided between adding a Debussy or a Beethoven to your collection. Hence, $P(B_1;B_2) = P(D;B_1) = P(D;B_2) = \frac{1}{2}$.” It then “follows readily that $P(D;B_1;B_2) = \frac{1}{3}$.” However, this conclusion “is unacceptable on *intuitive grounds* because the basic conflict between Debussy and Beethoven is not likely to be affected by the addition of another Beethoven recording” [Tversky (1972), p.283, emphasis added]. The empirical evidence had made the normative theory less useful for practical purposes, but it was this last introspective argument that dealt the final blow.

Thus, a new normative decision theory was required, Tversky argued, and this should preferably be a theory that could serve both the normative and the descriptive domain. Tversky then proposed such a theory, labeled elimination-by-aspects (EBA). Elimination-by-aspects was as simple and elegant as it was convincing. Rational human decision making, Tversky argued, occurs not through a process of expected utility maximization, but through a sequential process of eliminating the alternative with the lowest expected value. Here I quote Tversky at length.

The present development describes choice as a covert sequential elimination process. Suppose that each alternative consists of a set of aspects of characteristics, and that at every stage of the process, an aspect is selected (from those included in the available alternatives) with probability that is proportional to its weight. The selection of an aspect eliminates all the alternatives that do not include the selected aspects, and the process continues until a single alternative remains. If a selected aspect is included in all the available alternatives, no alternative is eliminated and a new aspect is selected. Consequently, aspects that are common to all the alternatives under consideration do not affect choice probabilities. Since the present theory describes choice as an elimination process governed by successive selection of aspects, it is called the elimination-by-aspects (EBA) model.

[Tversky (1972), p.285]

Tversky's EBA model is an example of attempts made by (behavioral) decision theorists in the late 1960s and early 1970s to move away from the traditional normative decision theory in order to propose alternatives. It shows that Tversky, over the course of the first eight years of his professional career, became increasingly dissatisfied with the normative theory as set out by Savage. Moreover, it shows that by the late 1960s he was actively searching for a solution in view of the difficulties, and that he sought to construct a new normative theory.

2.6 Caught between a priori axioms and behavioral deviations

Tversky was professionally trained at the University of Michigan during the 1960s in two related traditions: the mathematical psychology of Louis Leon Thurstone, Stanley Stevens, and Clyde Coombs, and the decision theory/behavioral decision research of Leonard Savage and Ward Edwards. In the work of Tversky these two branches of psychology came together, extending and influencing each other. The problem that arose was that the experiments produced systematic behavioral deviations that potentially disproved the very foundations of measurement theory and decision theory. Because of the far-reaching consequences of accepting that the axioms were falsified, Tversky was reluctant to accept this conclusion. At the same time, however, he took the behavioral deviations seriously and was unwilling to accept as an explanation for the behavioral deviations *ad hoc* explanations that problematized the

experimental procedure, such as the solution of changing tastes between experimental sessions. By the late 1960s and early 1970s, Tversky was actively looking for a way to solve the behavioral deviations problem. His eliminations-by-aspects theory was an attempt to take the behavioral deviations seriously, and to reconstruct decision theory and measurement theory based on a new foundation.

The methodological tension Tversky was struggling with involved how a theory that is *a priori* true can be combined with experimental results that point in many directions, but only occasionally in the direction of the *a priori* theory. Tversky could not proceed as an experimentalist who seeks to test whether a theory is right or wrong. As the axioms were *a priori* truths, they could only be proven wrong on the basis of *a priori* reasoning. Measurement theory and normative decision theory were simply not devised and employed as theories that could be proven wrong experimentally. At the same time, however, Tversky wanted to do justice to the experimental results he had obtained. Sticking to the axioms of measurement theory and decision theory would have implied that whenever a systematic behavioral deviation was observed in the experiments, there was something wrong with the experiment. This would mean that the majority of Tversky's experiments were invalid. Tversky was thus effectively caught between the *a priori* truth of the axioms of measurement theory and decision theory, and the behavioral deviations that surfaced in his experiments. He had to decide between either taking his experiments seriously, or accepting the axioms of measurement theory and decision theory. His elimination-by-aspects theory proves that by the early 1970s, Tversky had chosen the first option. He accepted his experimental results as valid and thus had to construct a whole new basis for measurement theory and decision theory. But he did so by disproving Savage's decision theory on intuitive grounds. That is, by accepting that ultimately only an *a priori* 'test' of the axioms could prove them wrong.

For reasons that will become clear in Chapter four, the elimination-by-aspects theory turned out to be a road not taken. In the beginning of the 1970s, Kahneman offered Tversky a solution that both solved the problem of experimental behavioral deviations, and at the same time left intact the fundamentals of measurement theory and decision theory. The Savage-Edwards approach to decision theory and behavioral decision research continued to be problematic for Tversky, but the solution he would come to favor lay in another direction. Tversky would find this solution in his joint work with Kahneman.

3. Kahneman's cognitive mistakes

3.1 From correlational to experimental psychology

Kahneman obtained a B.A. from Hebrew University in 1956 while working as a psychologist in the Israeli army. In 1958 he moved to San Francisco and obtained a PhD from the University of California at Berkeley in 1961 under the supervision of Susan Ervin (1927-). To understand how Kahneman solved Tversky's problems in mathematical psychology, decision theory and behavioral decision theory in the 1970s, we need to examine the research Kahneman conducted before his collaboration with Tversky. Between 1961 and 1971, Kahneman's research was about semantic differentials, optometry, vision research, and related themes. Kahneman's research in this period was unrelated to mathematical psychology, unrelated to decision theory, unrelated to behavioral decision research, and despite some retrospective hints of Kahneman to the contrary, entirely unrelated to economics.

Based on Kahneman's recollections in his autobiography and the one publication that emerged from this, his early work for the Israeli army in the early 1950s and at the Hebrew University is best characterized as correlational psychology [Danziger (1990,1997), Gigerenzer (1987a,b)].²¹ Correlational psychology builds theories on the basis of correlations in statistical data, for example between IQ and the degree of education. Using methods developed by the British army in World War Two, the aim of Kahneman's early research was to develop reliable predictions about the future performance of people on the basis of character traits, be it in the army or in different kinds of jobs. For instance, to find out at an early stage which new recruits in the army would eventually be successful leaders on the future battlefield, different tests were designed to evaluate the differences between recruits with respect to a few behavioral and personal characteristics that were thought to relate to leadership capacities.

It is not difficult to see that in this kind of research the ability of the researcher to predict the future performance of the subjects investigated is an important, and perhaps the only way to measure success. A classification of new recruits in the army along different dimensions might be an interesting exercise, but if it does not predict

²¹ Kahneman and Ghiselli (1962), Kahneman (2002). After he completed his PhD at Berkeley in 1961 Kahneman returned to the psychology department at Hebrew University where he would remain until 1978. In the meantime, however, he was a visiting scholar at the University of Michigan in 1965-1966, a visiting scientist and lecturer in psychology at Harvard University in 1966-1967, a visiting scientist during the summer terms of 1968 and 1969 at the Applied Psychology Research Unit of Cambridge, UK, and a lecturer in the graduate program of the University of Michigan in 1968-1969.

better than chance then it is of no use. In his autobiography [Kahneman (2002)] Kahneman recalls how frustrating it was when time and again he was confronted with the fact that his predictions were anything but reliable. Extensive questionnaires and tests were set up, but in the end it turned out that the intuitive guesses of the staff members who conducted the tests and collected the questionnaires proved better than the scientific predictions.

Dissatisfied with the results of this research and eager to develop his research skills, Kahneman switched to experimental psychology of vision, resulting in some twenty-five articles over a period of ten years, including two publications in *Science*, and a whole range more in prominent experimental psychology journals such as the *Journal of Experimental Psychology*. There is not one particular theme or article that stands out during the decade from 1961 to 1971. The psychological view held by Kahneman emerges when the different themes and articles are considered next to each other. In the following two sections I will first provide an overview of Kahneman's work during this period and then come to a general assessment of Kahneman's work in this period.

3.2 Kahneman's research 1961-1971

Kahneman's career began with the theoretical work he did concerning the models used in experimental studies of semantic differentials. Semantic Differential (SD) research investigates people's attitude towards words [Heise (1970), p.235], or, put differently, measures the meaning of abstract objects to the individual [Kiddler (1981)]. A distinction is made between the denotative and the connotative meaning of a word, or concept, or object. Thus, it is assumed that apart from the dictionary, or the denotative meaning, words are also assigned connotative meanings by the individual. The words *massacre* and *rape*, for instance, are attributed different connotations than *flower* or *sunny day* in terms of good versus bad. SDs are measured on a bipolar scale, for which in principle all opposites can be used, such as good-bad, soft-hard, fast-slow, clean-dirty, valuable-worthless, and so forth.

Articles on SD make up a small, but important part of the publications of Kahneman's early work. Apart from his dissertation, it was the subject of one published article.²² It also provides a good illustration of Kahneman's take on

²² Kahneman's dissertation consisted of a paper he wrote in eight days [Kahneman (2002), p.6].

experimental psychology. Within the field of SD research, Kahneman's focus was on the theory behind the models that are used to infer conclusions about the connotative meanings. In Kahneman (1963) he showed that models that are used to measure SDs are mathematically not sufficiently sophisticated and may give rise to wrong interpretations of what is observed. Kahneman considered the following model for the rating s_{ijk} of the concept j by individual i on scale k

$$s_{ijk} = T_{jk} + C_{ik} + d_{ijk} \quad [1]$$

in which T_{jk} is the “true score” of concept j on scale k , computed as the average score of a number of judges. C_{ik} is the “constant deviation” of subject i on scale k , computed as the average deviation of subject i from the true score over a large number of concepts. d_{ijk} is the “specific deviation” (or “error of judgment”) on a particular rating. Kahneman argued that the practice of contemporary SD research wrongly assumed that the specific deviations of ratings were uncorrelated. For example, deviating from the true score could very well be correlated during the course of one experiment. Improvement should be sought in the direction of more “precise algebraic” models. Kahneman's research focused not so much on the theory of SDs as such, but on the improvement of the analysis of variance in the statistical models it employed. Specifically, he focused on the notion of the “error of judgment” in SD research. The object of investigation was to understand how people deviate from what is true or correct. In effect, that meant that the analysis of the actual process of how people make judgments was black-boxed.

In 1962-1963, Kahneman set up a vision lab at the department of psychology of the Hebrew University [Kahneman (2002), p.6]. Many of the articles he published in the following years were derived from the experimental results of this lab. In this research, Kahneman investigated the relationship between the “energy” of different stimuli and visual perception capacities. “Energy” was employed as a general concept to define the strength of a stimulus; the brighter, the more illuminated, the more contrasted, the longer and so forth the stimulus was, the more energy it had. Visual perception was measured in terms of the reaction times of the subjects. In the typical experiment, the subject had to decide as quickly as possible whether the opening of a

so-called Landolt C was directed up-, down-, left-, or rightwards.²³ The conditions in terms of brightness, contrast, and so on in this setting could be varied in numerous ways. The visual task could also be combined with other cognitive tasks. Kahneman's text book on the psychology of vision and attention, *Attention and Effort* (1973), is still used in the early twenty-first century as standard reference on the subject [Dawes – Interview (2008)].

Examples of this research include Kahneman's (1964), "Control of Spurious Association and the Reliability of the Controlled Variable" and Kahneman's (1966b), "Time-Intensity Reciprocity in Acuity as a Function of Luminance and Figure-Ground Contrast."²⁴ In Kahneman and Norman (1964), the relation between the minimal amount of time subjects need to identify a visual stimulus (labeled the "critical duration" t_c) and the energy in terms of brightness and duration of the stimulus is investigated. It was shown that the stimuli of equal energy do not necessarily produce the same critical duration and that a given visual stimulus does not trigger one but multiple sensory processes. The second conclusion particularly opposed the general view held in the psychophysical community that one stimulus triggers only one sensory process. In Kahneman (1966b) the Bunsen-Roscoe law of the time-intensity of reciprocity is (partly) falsified. This law forms a central concept in the psychophysics of vision, stating that up to the critical duration t_c , duration and intensity of the stimulus are interchangeable, in which duration is measured in seconds and intensity in lux. Kahneman showed that time-intensity reciprocity fails to hold when a Landolt C stimulus at 40mL*msec is preceded by a 2 second flash of 1mL.

In the psychophysical paradigm, visual perception is seen as one of many cognitive tasks. Other cognitive tasks include conversation, or more generally, speech, learning, and calculation. How different cognitive tasks influence one another was investigated in Kahneman and Beatty (1966, 1967), Kahneman et al. (1967,1968), and Kahneman and Peavler (1969). The explicit emphasis in these articles was on how the combination of different cognitive tasks could lead to "errors of judgment." In Kahneman et al. (1967) for instance it was shown that the capacity to visually

²³ The Landolt C is one of the standard symbols used in American psychophysics of vision and optometry. It consists of a C in which the opening can be varied, and which is either surrounded by bars the width of which equals the C's opening or not surrounded.

²⁴ Other examples include Flom et al. (1963), Kahneman (1965a,b, 1966a, 1967), Kahneman and Norman (1964), and Kahneman et al. (1967).

perceive substantially decreases when subjects were engaged in other mental tasks such as speech or calculation. The “error of judgment” in these cases is very real, as it explains for instance why car drivers may miss a stop sign when engaged in conversation. It again illustrates Kahneman’s focus on the psychology of mistakes.

Another way in which the psychophysics of vision and cognitive tasks are connected is through measuring a mental task. For example, a clear correlation can be found between the difficulty of the mental task and the diameter of the pupil. The pupil dilates when the task begins and constricts when the answer or report is given. This relation is investigated in Kahneman and Beatty (1966), Kahneman et al. (1968), Kahneman et al. (1969), and Kahneman and Wright (1971). In Kahneman and Wright (1971), the authors investigated the correlation between pupil size and short-term memory. It was shown that pupillary size provides a good measurement for mental activity. When involved in a mental task, the subjects’ pupil clearly dilates. A number of characteristics of short term memory were reported. Mental activity in case of short term memory seemed for example negatively correlated with the time subjects have to store a stimulus in memory. If a stimulus has to be stored in memory for seven seconds, instead of three, mental activity on average was lower. However, there are more factors influencing the level of activity. It was also shown that mental activity and the difficulty of a memory task are positively correlated. Kahneman et al. (1969) compared the pupil diameter as a measurement for mental activity with measurements of the heart rate and of skin resistance. The results indicated that these three measurements are not very well correlated. Pupil diameter remained therefore the preferred method. Nevertheless, some reservation was expressed, as the relation between mental activity and each of the three methods is not precisely known.

3.3 Kahneman’s cognitive errors

In Kahneman’s vision research an emphasis was placed on the question under which circumstances the human mind makes cognitive errors. Kahneman showed that there is a trade-off between different cognitive tasks in perception capacities, and that as a result people may sometimes “fail” to perceive the stimulus and make an error in judgment. Furthermore, the research done by Kahneman in the period between 1961 and 1971 was in line with the interwar drive to eliminate all introspection from psychology [Danziger (1997)]. In Kahneman’s experiments self-reports were not necessary to establish how the cognitive system operates. The behavior of the

cognitive system could be inferred from observed behavior and physical responses which cannot be controlled, such as pupil dilations and restriction. The human mind was considered to be a black box whose functioning could be inferred from the uncontrollable and unconscious responses made by the individual subjects.

Both elements are important in gaining an understanding of Kahneman's psychology and his subsequent influence on Tversky. The recurring theme of the cognitive errors shows, that in Kahneman's view, psychology was about discovering how people deviate from a norm behavior. This aspect of experimental psychology dates back to the beginning of experimental psychology in nineteenth-century Germany. But in nineteenth-century German and interwar American experimental psychology, this framework was adopted for the purpose of discovering what the *true* value was. The experimental psychologists wanted to know the true value of, for instance, the smallest amount of difference in weight people could perceive, and for this purpose devised a framework, which in spite of all the individual errors, could establish the true value. Thurstone, for instance, wanted to measure *the* attitude towards religion of a group of people, and for this purpose he constructed a method that would elicit the attitude from a series of observations in which each individually deviated from the true value. Experimental psychology was explicitly modeled after experimental practice in physics, where the physicist tries to establish the true value of boiling water by conducting a series of measurements in which each measurement individually deviates from the true value and from each other.

Kahneman employed the experimental psychological framework, but applied it differently than the nineteenth-century German psychologists. In Kahneman's work the true value was known. The true value was an accurate prediction of a recruit's future leadership capacities, or the true value was not running through a traffic light when driving a car. The question Kahneman then raised was how, when, and why the cognitive machinery fails to act according to the true value. Kahneman used an experimental psychological framework, but applied it with the opposite purpose. He did not want to find out what the *true* value was, but how people *deviate* from the true value. In Kahneman's research, the true value was always clear and determined by the experimenter. Kahneman knew how the cognitive machinery ideally responds, and investigated whether it actually does so. In Kahneman's understanding, the scientist thus completely determined in each experimental situation what the good, optimal, or rational behavior should be. This was in line with the scientific desire to

eliminate all introspection because it assumed that the experimental subject cannot judge whether he or she is giving the correct response or not. In Kahneman's experiments the experimenter determined how the subject should behave and determined how it did behave. All authority for judging behavior was placed in the hands of the scientist.

Because Kahneman has never provided an extensive theoretical exposition of the assumption that human beings often make cognitive errors, one could easily dismiss it as merely a nice way of illustrating theories which are perhaps not too exciting, but that would be a mistake. The key to understanding Kahneman's psychology lies in his conviction that human beings often make cognitive errors. Since Kahneman is a psychologist who sets out his theories through case-based reasoning, it is also by means of these cases that we can best illustrate and understand Kahneman's firm belief in cognitive errors. In the example of the army psychologist at the beginning of this chapter, Kahneman and his colleagues really believed that through their extensive studies they could accurately predict, or at least predict better than by mere chance, the future performance of different candidates for a job. The fact that they could not was for the young Kahneman a true cognitive illusion that he needed to correct for himself [Kahneman (2002)].

Another illustrative example recalled by Kahneman in his autobiography was the moment a flight instructor disagreed with the psychologists' theory that praise is more effective in developing skills than punishment. The flight instructor reasoned that although he praised the good performance of his recruits, the next time the performance would almost always be worse. Similarly, he would always punish recruits who had done a poor job, and this would almost always improve performance the next time. To Kahneman this was a clear cognitive illusion. A good performance is statistically more likely to be followed by a worse performance than by an equally good or even better performance, and vice versa. Thus, the flight instructor was suffering from a cognitive illusion. The truck or car driver described above who is engaged in a conversation and thus does not see a traffic light that he or she would otherwise not miss, really does make an error. His or her cognitive apparatus is tuned to noticing traffic lights, but it fails to do so.

To Kahneman it was and is a given fact of life that human beings often make cognitive errors. However, science can help in two ways. First, scientists can set out what the correct way of behaving should be for each situation. For the truck driver, it

is obvious what the correct behavior is, but for the flight instructor it may not be intuitively clear what the correct way of reasoning is. Scientists can therefore help to establish the correct way of reasoning. Second, scientists, and in particular psychologists, can help by investigating when, how and in what way human beings make cognitive errors and thus provide a basis for designing tools or education to help human beings correct these cognitive errors.

4. Conclusion

Before Kahneman and Tversky started collaborating in 1969 and published their first paper in 1971, the research that each conducted was not directly related. Although both can, partly in the case of Tversky, be placed in the realm of experimental psychology, the approach each took and the psychological phenomena they investigated differed. Tversky was raised in the fields of mathematical psychology, decision theory, and behavioral decision research at the University of Michigan. He was principally interested in human decision making under risk. Kahneman, on the other hand, conducted research on semantic differentials, vision and the interaction of cognitive tasks.

The point of contact between their two research programs was that of experimentally observed behavior that deviates from what the scientist expects. Tversky was struggling to incorporate behavioral deviations in the representational theory of measurement and decision theory, whereas Kahneman was investigating how, when, and why human beings deviate from the response or behavior they should display. This concern with deviating behavior would become the basis for their collaborative research of the 1970s.

4. Heuristics and Biases for psychology and economics, Kahneman and Tversky in the 1970s

1. The Kahneman and Tversky collaboration

In 1969 Daniel Kahneman and Amos Tversky started a collaboration that would result in twenty-one collaborative papers and two co-edited books, including one published together with Paul Slovic. They continued to co-operate on different projects until Tversky's death in 1996, but the most productive and creative period was from 1969 to 1979, including the widely cited 1974 *Science* and 1979 *Econometrica* articles. The cooperation was initiated by Kahneman, who was looking for ways to experimentally test his intuition that an individual's cognitive apparatus often fails, and who tried to find a theory that might account for these cognitive errors [Krantz – interview (2008), Dawes – interview (2008), Kahneman (2002)]. But the ensuing program stemmed as much from the result of Tversky's growing doubts concerning Leonard Savage and Ward Edwards' assumption that, generally speaking, individuals decide according to the normative rules of logic, Bayesian statistics, and expected utility theory.

Although their collaborative work constituted an important part of their research, especially in the 1970s, it was never the only project they were engaged in. Kahneman continued to work on vision research and Tversky kept working on measurement theory and, sometimes, on elimination-by-aspects. The role that the collaborative work with Kahneman in Tversky's life played is nicely illustrated in fifteen letters which Tversky wrote to his close friend David Krantz between 1967 and 1977. His work with Kahneman is briefly mentioned for the first time in 1969, when Tversky, in an off-hand remark, notes that "I am working a little bit with Danny, on the problem of statistical intuition, which helped to reinforce my prejudices concerning the importance of statistics" [Tversky's letter to Krantz, October 5, 1969]. A month later, the work with Kahneman seems to have taken off seriously, as one project among a number of different projects Tversky was working on:

Danny and I got deeply involved in the problem of processing uncertainty: we are running a research seminar and a couple of studies on the topic [...] I am working now on the Chapters of our book [*Foundations of Measurement I – FH*]. The editing apparently takes much more time than I realized, but is certainly worth doing. What has happened to our MDS paper submitted to JMP?²⁵ [Tversky's letter to Krantz, November 2, 1969].

When two years later he wrote about the re-organization of the psychology department he and Kahneman were seeking to advance, Tversky briefly mentioned what would become the famous 1974 *Science* publication. “Danny and I are writing a sort of review paper on our work for someplace like *Science*, and I returned to Chapter 16 [of *Foundations of Measurement II – FH*]” [Tversky's letter to Krantz, November 14 1971].

The foundation for their collaborative fame in psychology, as well as the basis for their influence on economics from the early 1980s onwards, was laid in the 1970s. Between 1971 and 1979 Kahneman and Tversky co-authored eight articles. The first seven articles form part of the Heuristics and Biases approach, a new approach in behavioral decision research developed in the early 1970s. The prospect theory paper, published in 1979 in *Econometrica*, further developed the Heuristics and Biases program and was aimed explicitly at entering and influencing the economists' debate on individual human behavior. In what follows, I will first set out the Heuristics and Biases approach, and show in what way this approach was a mix of the earlier work done by Kahneman and Tversky. After this, I will set out prospect theory and explain how this was intended to plead an argument in economics. The conclusion infers what consequences Kahneman and Tversky's theories had for the conception of rationality in psychology and economics.

2. Heuristics and Biases

When in 1969 Kahneman and Tversky started to cooperate, their joint work became a mix of their earlier individual research. Tversky's work on decision theory, with its distinction between the normative and descriptive realm, became coupled with

²⁵ “MDS paper” and JMP referred to Tversky and Krantz (1970) “The dimensional representation and the metric structure of similarity data,” published in the *Journal of Mathematical Psychology* (JMP) [email Krantz to author August 11, 2008].

Kahneman's psychology of mistakes. The first Heuristics and Biases articles showed that human beings in the real world display behavior that in all kinds of ways systematically deviates from what is normatively correct.

For their first article Tversky posed a set of questions to eighty-four participants who attended the 1969 meetings of the American Psychological Association and the Mathematical Psychology Group that meant to capture Kahneman's personal experience of incorrect research planning and unsuccessful replications, as discussed in Chapter three. "Suppose," Kahneman and Tversky asked, "you have run an experiment on 20 Ss, and have obtained a significant result which confirms your theory ($z = 2.23$, $p < .05$, two-tailed). You now have cause to run an additional group of 10 Ss. What do you think the probability is that the results will be significant, by a one-tailed test, separately for this group?" [Kahneman and Tversky (1972), p.433]. The answer depends on the interpretation of the information provided. However, it should be below but close to 0.50, Kahneman and Tversky argued. Nine out of the eighty-four participants gave answers between 0.4 and 0.6, which Kahneman and Tversky interpreted as "reasonable." The other seventy-five, however, gave answers that exceeded 0.60. The median response of all participants was as high as 0.85. Thus, even those professionals who were trained and who were explicitly asked to give the normatively correct answer failed to calculate it correctly. Kahneman and Tversky felt justified in inferring the strong and bold thesis "that people have strong intuitions about random sampling; that these intuitions are wrong in fundamental respects; that these intuitions are shared by naïve subjects and by trained scientists; and that they are applied with unfortunate consequences in the course of scientific inquiry" [Tversky and Kahneman (1971), p.105].

Kahneman and Tversky found it appalling and fundamentally disturbing to see that even trained professionals failed to behave according to the dictates of normative theory. Why did the majority of them fail? As set out in Chapter three, Tversky's answer earlier would have been that either there had been something wrong with the experiment, or that the normative theory was wrong. This time, however, Kahneman and Tversky took a different route. Building on the work of William Estes (1919-), a mathematical psychologist renowned for work on learning theory he conducted while at the University of Minnesota during the 1940s-1960s [e.g. Estes (1964), Bower (1994)], Kahneman and Tversky hypothesized that individuals have the tendency to suppose that a sample from a population must represent the population in its general

characteristics. In other words, they accounted for their results by supposing that human nature makes individuals ignore the possibility that a sample of a population may not be an accurate representation of that population. Kahneman and Tversky hypothesized that human nature sometimes provides individuals with the wrong intuition and that as a result they fail to give the right answer. However, Kahneman and Tversky took the research of Estes a step further by concluding that if one considers a sample to be representative of its population, then it could be thought of as a “heuristic.” They advanced the idea the human mind uses this heuristic to base decisions on.

It is not clear where Kahneman and Tversky derived the term ‘heuristic’ from. It appeared for the first time in 1971 without any precursors in either Kahneman’s or Tversky’s earlier work, and from the beginning it was used as a natural term for an intuitive response. The same term was used by Herbert Simon [e.g. Heukelom (2007)], leading one to suppose that Kahneman and Tversky had obtained the term from him. But Simon used the term differently (on which more below in section 4.3) and is not mentioned in Kahneman and Tversky’s research of the early 1970s. Thus, it seems that the term ‘heuristic’ was a general term that psychologists could use to refer to intuitive, automatic behavior of individuals [Krantz – interview (2008), Dawes – interview (2008)].²⁶

The reason, according to Kahneman and Tversky, why the majority of scientists and lay persons systematically deviated from the norm-answer that was given in Tversky and Kahneman, “Belief in the Law of Small Numbers” (1971), and further developed in Kahneman and Tversky, “Subjective Probability: A judgment of representativeness” (1972) was that human beings, in general, do not base their decisions on the normative laws of, in this case, probability theory and statistics, but instead use a “representative heuristic.” Kahneman and Tversky described the representative heuristic as the phenomenon that “[t]he subjective probability of an event, or a sample, is determined by the degree to which it: (i) is similar in essential characteristics to its parent population; and (ii) reflects the salient features of the process by which it is generated” [Kahneman and Tversky (1972), p.430]. As a result of this representative heuristic, most of the professional psychologists mentioned in

²⁶ The use of the term ‘heuristic’ needs more investigation. Note in this regard that also the term ‘behavioral economics’ was used by Kahneman and Tversky and their followers from the 1990s onwards without (hardly) any reference to the already existing behavioral economic program in the Herbert Simon and George Katona traditions.

the example estimated the probability requested to be much higher than it actually was (the median estimate was 0.85). Because human beings have much more faith in small samples than they should, Kahneman and Tversky half jokingly labeled this phenomenon the “belief in the law of small numbers,” in reference to the law of large numbers. The analogy with faith and belief expresses Kahneman and Tversky’s view that an individual’s erroneous behavior is the result of false beliefs for which the unenlightened individual cannot really be blamed. The “deviations of subjective from objective probability seem reliable, systematic, and difficult to eliminate” [Kahneman and Tversky (1972), p.431], and “[t]he true believer in the law of small numbers commits his multitude of sins against the logic of statistical inference in good faith.”²⁷ The representation hypothesis describes a cognitive or perceptual bias, which operates regardless of motivational factors” [Tversky and Kahneman (1971), p.109].

3. Kahneman’s case-based reasoning

Kahneman and Tversky employed a case-based reasoning that finds its origin in Kahneman’s research of the 1960s. Kahneman’s research on the semantic differential, in particular, was never far away. For instance, the reason that people’s judgments systematically deviated from the correct solution was because of the individual’s connotation of the event of which the probability was to be judged: “Although the “true” probability of a unique event is unknowable, the reliance on heuristics such as availability or representativeness, biases subjective probabilities in knowable ways” [Tversky and Kahneman (1973), p.231]. The way in which an individual’s connotation of words in semantic differentials research systematically deviated from the average, and thus “true” connotation, was the same way in which the individual’s connotation of the probability of events systematically deviated from the objective, and thus “true” probability of those events. Given that framework, different explanations in terms of fixed cognitive rules could and were then put forth.

Typically, the argument in Kahneman and Tversky’s research was made not so much by giving theoretical explanations for why such and such was a good theory or account of observed behavior, but by supplying examples of the experimental

²⁷ The notion of subjective probability does not always have the same meaning. In Savage (1954), for instance, it is the rationally calculated probability of and by the individual. In Tversky and Kahneman, it is the subjective perception of objective probability.

questions subjects had been asked which were meant to give the reader an intuitive understanding of the point they were trying to make. Similarly, Kahneman and Tversky's adversaries have often proceeded by deconstructing their examples and illustrations, or by giving counter-examples. In the typical counter argument it was shown that a different conclusion could be inferred from the observed behavior, or that responses to another set of hypothetical questions falsified Kahneman and Tversky's conclusions [e.g. Gigerenzer (1991,1993,1996), Hertwig and Gigerenzer (1999), Lopes (1991)]. For a good understanding of Kahneman and Tversky's work one needs first to have a feeling for the kind of questions they asked, the examples they gave, and how they inferred general conclusions from them. Therefore, I will briefly introduce and discuss Kahneman and Tversky's most frequently used examples and illustrations.

Assuming that the probability of a new-born to be a boy or a girl is 0.5, consider the following question.

All families of six children in a city were surveyed. In 72 families the *exact order* of births of boys and girls was G B G B B G.

What is your estimate of the number of families surveyed in which the *exact order* of births was B G B B B B? [Kahneman and Tversky (1972), p.432, emphasis in the original]

The normatively correct answer is 72, as any sequence of boys and girls is equally probable. However, average estimates of the second sequence were systematically lower than the first. People, in other words, incorrectly believed the first sequence to be more probable than the second, from which Kahneman and Tversky concluded that “[p]eople view chance as unpredictable but essentially fair” [Kahneman and Tversky (1972), p.435]. An alternative, but related explanation was that people judged the first sequence to be more probable than the second because it better represented their image of a family with six children (representativeness), or that an image of a family with three boys and three girls was more readily available than an image of a family with five boys and one girl (availability).

In another experiment to test human's capacity to reason probabilistically, subjects were posed the following question.

A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

- (i) 85% of the cabs in the city are Green and 15% are Blue.
- (ii) A witness identified the cab as a Blue cab. The court tested his ability to identify cabs under the appropriate visibility conditions. When presented with a sample of cabs (half of which were Blue and half of which were Green) the witness made correct identifications in 80% of the cases and erred in 20% of the cases

Question: What is the probability that the cab involved in the accident was Blue rather than Green? [Tversky and Kahneman (1980), p.62]

The majority responded 80 percent, which was probably based on how often the witness had identified the color correctly. However, again they failed to take into account the base-rate distribution. Using Bayes' theorem the normatively correct answer is just over 41%.

The most well-known question amongst the many experimental questions of Kahneman and Tversky has become the so-called 'Linda problem,' no less because it has often been used by Kahneman and Tversky's adversaries. The Linda problem is as follows:

Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Which of the following two alternatives is more probable:

- 1) Linda is a bank teller
- 2) Linda is a bank teller and active in the feminist movement. [Tversky and Kahneman (1983), summarized from pp.297 and 299]

On average, there was a strong bias towards judging 2) to be more probable than 1), this despite the fact that 2) is logically contained in 1). Because this bias is an

illustration of the failure to see that the probability of the conjunction of two or more events can never exceed the probability of one of the events, this bias was labeled the conjunction fallacy.

From 1972 to 1974, on the basis of these and similar examples, Kahneman and Tversky developed the conclusion that their experimental evidence contradicted Edwards' behavioral decision research view that despite some unresolved issues, the normative models generally worked well descriptively. Extending further Tversky's earlier work, Kahneman and Tversky argued that the normative models worked descriptively far worse than had previously been thought. But instead of arguing that either the experiments were flawed or that the normative theory was wrong, the two options available in Tversky's research on human decision making in the 1960s, they now argued that the experimental results were perfectly valid and that there was nothing wrong with the normative theory. Instead, the new conclusion that was drawn was that when individuals make their decisions intuitively, they systematically deviate from the rational norm. The argument was thus directed explicitly against Tversky's former mentor and collaborator. With respect to Bayes' rule, Edwards had mistakenly assumed "that man, by and large, follows the correct Bayesian rule, but fails to appreciate the full impact of evidence, and is therefore conservative" [Kahneman and Tversky (1972), p.43]. However, the mainstream representatives of signal detection theory were also criticized.²⁸ "Peterson and Beach (1967), for example, concluded that the normative model provides a good first approximation to the behavior of the Ss who are 'influenced by appropriate variables and in appropriate directions'" [Kahneman and Tversky (1972), p.43]. Kahneman and Tversky had come to fundamentally disagree with them. "[In] his evaluation of evidence, man is apparently not a conservative Bayesian: he is not Bayesian at all" [Kahneman and Tversky (1972), p.449].

The alternative theory Kahneman and Tversky proposed was their Heuristics and Biases theory, first labeled as such in Tversky and Kahneman (1974), "Judgment under Uncertainty: Heuristics and Biases." In this theory, people do not use the normative theories of probability and logic to make decisions under uncertainty, but instead rely on a number of heuristics, heuristics that sometimes lead to systematic

²⁸ As said above in Chapter two, Signal detection theory (SDT) is a branch of psychophysics that investigates the individual's ability to distinguish between signal and noise. In other words, it investigates decision making under noisy conditions. See e.g. Green and Swets (1964).

deviations. In the often quoted definition of the theory, *Heuristics and Biases* “shows that people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors” [Tversky and Kahneman (1974), p.1124]. Kahneman and Tversky emphasized the importance and functioning of a few heuristics, such as representativeness, availability, and anchoring. But by no means was the *Heuristics and Biases* theory meant to remain confined to these few heuristics. There was no limit to the number of heuristics that possibly could be discovered in humans’ minds. The *Heuristics and Biases* program summed up the many violations of the normative models Kahneman and Tversky had found, and provided a small, non-exhaustive list of explanations that might account for these violations.

4. How to understand Heuristics and Biases

4.1 The collaboration

Kahneman and Tversky’s collaborative work in *Heuristics and Biases* was a combination of the research conducted by them in the 1960s. *Heuristics and Biases* was closely related to Tversky’s earlier work. It was about human decision making and referred to much of the same literature as Tversky’s work in the 1960s. Moreover, the psychological community considered *Heuristics and Biases* to be a part of behavioral decision research [Phillips and Von Winterfeldt (2006), p.8]. The link with Kahneman’s earlier work is less obvious; *Heuristics and Biases* had little to do directly with vision and attention research, optometry, semantic differentials or personnel psychology. But Kahneman had a profound influence on a conceptual level; the adjustments of Edwards’ behavioral decision research made by Kahneman and Tversky were the result of the psychological framework developed by Kahneman in the 1960s.

In Chapter three I showed how towards the late 1960s and early 1970s Tversky became increasingly dissatisfied with the approach and theory of behavioral decision research and decision theory. The normative models were consistently violated by subjects, and there did not seem to be an explanation for this. Elimination-by-aspects was an attempt to solve this problem by providing a new normative theory of rational decision making. However, its merits had not yet been tested and its implications for measurement theory were not clear. Kahneman suggested an

alternative route by introducing the idea that in every decision situation there is but one optimal, or normative, solution. To Kahneman it did not make sense to test a number of different normative models to see which, if any, fit best. When a decision problem under uncertainty occurred, there was always only one normative solution and that was the solution determined by logic, Bayesian statistics, and expected utility theory.

Kahneman emphasized that there was absolutely no reason to doubt these normative rules, as Tversky had become inclined to do. Irrespective of whatever people think of the norms or in whatever way they behave in daily practice, the normative rules of logic, Bayesian statistics and expected utility theory were the fixed rules of rational behavior. Kahneman proposed the view that when people were observed to violate the normative rules, this meant that they had made an error, a mistake. Kahneman, in other words, introduced his psychology-of-mistakes view to Tversky's behavioral decision research. If people were observed to violate the normative rules, this meant that people made consistent and systematic errors.

Contrary to Tversky's earlier work and contrary to the framework as set out by Edwards and Savage, experimental violations no longer implied there might be a problem concerning the normative rules. Thus, Kahneman saved the normative basis of decision theory and measurement theory. Instead of placing the burden on the theory, as was done by Savage and Edwards, Kahneman proposed that the burden of the violations be placed on the human beings in the experiments. *Heuristics and Biases* was a research program in behavioral decision research that built directly upon the earlier work done by Tversky in the 1960s, and still drew on the same authoritative sources. However, after it became inspired by Kahneman's work of the 1960s, the approach had changed conceptually in fundamental ways.

Heuristics and Biases took Tversky in a very different direction than his elimination-by-aspects theory. In elimination-by-aspects Tversky had started from the position that the experimental results were valid and that, by and large, the normative theory should account for observed behavior. Given Savage's assumption that all rational people should behave in accordance with the normative theory, Tversky concluded that Savage's normative set of axioms must be wrong and set himself to developing a new normative theory. In *Heuristics and Biases*, however, Kahneman and Tversky departed from Savage and Edwards' starting point in a different way. They reasoned that Savage's normative axioms for rational decision making were

valid irrespective of what decisions human beings actually made. They now concluded that the experimental results were valid, but that at the same time Savage's normative rules were the one and only set of rules for rational decision making. As a consequence, behavior that deviated had to be considered an erroneous deviation from the rational norm. Kahneman pulled Tversky in a different direction than Tversky had initially chosen. That said, *Heuristics and Biases* was a psychological program that Kahneman and Tversky developed together. At no point during his career did Tversky's deep commitment to his joint research with Kahneman waver. Neither did he seriously continue to develop elimination-by-aspects or other accounts of human decision making. *Heuristics and Biases* was a joint product and enjoyed the continued support of both its authors.

4.2 Kahneman and Tversky's experiments

In their collaborative research during the 1970s, Kahneman and Tversky conducted different kinds of experiments than the two had done individually in the 1960s. In their 1960s research the experiments were done in laboratories, or laboratory-like settings. That was in line with received experimental practice in psychology. From the early 1970s onwards, however, the experiments consisted of questionnaires with hypothetical questions which mainly students were asked to fill out. These questionnaires could be distributed anywhere, to participants at conferences, to students during a course at the university, and in a shopping mall on a Saturday afternoon.²⁹ These questionnaires consisted of hypothetical questions and were cheap and easy to conduct.

Laboratory experiments required a setting in which all of the variables could be carefully monitored. However, Kahneman and Tversky's questionnaire experiments required only copies of the questionnaires and pens to fill them out. No separate laboratory space was needed and no payment of the experimental subjects was required, filling out the questionnaires took a few minutes at the most. The advantages of conducting questionnaire experiments compared to using previous experimental psychological methods were clear, but Kahneman and Tversky's departure from received practice in experimental psychology required a rationale. In 1979, Kahneman and Tversky defended their new method of experimentation and

²⁹ It is outside the realm of this dissertation to examine to what extent this new type of experiment was part of a more general development in psychology.

contrasted it with two other methods of investigation. First they distinguished the possibility of “field studies,” which used “naturalistic or statistical observation.” Evidently, this category referred to the practice of correlational psychology, such as Kahneman’s experiments for the Israeli army, although Kahneman and Tversky did not use this term. Field studies, according to Kahneman and Tversky, could yield important insights when a new field of research was opened, but in the end they could only provide “crude tests of qualitative predictions, because probabilities and utilities cannot be adequately measured” [Kahneman and Tversky (1979), p.265].

Secondly, they recognized the method of laboratory experiments, a reference to experimental psychological practice although, again, Kahneman and Tversky did not call it as such. Despite all of its advantages in the particular case of decision theory, laboratory experiments had the disadvantage that stakes could only be relatively small. In a laboratory experiment with real pay-offs it was for financial reasons difficult to conduct an experiment in which the subjects were asked to choose between, say, \$300 for certain or a 0.8 chance at obtaining \$400. A set of hypothetical questions did not have to reckon with this constraint. Laboratory experiments in psychology had furthermore been set up to measure probabilities and utilities as precisely as possible and therefore often large sequences of very similar decision problems were required. This characteristic feature of laboratory experiments, Kahneman and Tversky argued, made it questionable whether the results obtained in the laboratory could be related to behavior in the real world. It complicated “the interpretation of the results and restricts their generality” [Kahneman and Tversky (1979), p.265]. The third method, that proved to be the best for solving the problem they were investigating was “the method of hypothetical choices.” It solved both the immeasurability problem of field studies, and the external validity problem of laboratory experiments. It did however rely on the assumption that people “know how they would behave in actual situations of choice” and that they “have no special reason to disguise their true preferences” [Kahneman and Tversky (1979), p.265].

Tversky’s experiments conducted during the 1960s were in the ‘small world’- ‘grand world’ setting, an experimental requirement Edwards had obtained from Savage. In this setting, the crucial, but only requirement for experiments was that both the utilities and probabilities should be unambiguously clear to both the experimenter and the subject. As earlier set out in Chapter two, this entails that a designated space qualifies as a viable laboratory environment when the utilities and probabilities have

been defined and when it is clear to both experimenter and subject that each have the same figures in their minds, and thus that the subjects have been convinced they are not being deceived. This could, for instance, be done by tossing the dice in front of the subject and by showing that the possible pay-off (in the form of money, cigarettes or anything else) was at hand and could be offered immediately. When these requirements were fulfilled, experiments could be conducted everywhere: in prisons, in classrooms or in specifically designed laboratories.

The Savage-Edwards experiments were often long and repetitive in order to allow for initial adjustment behavior, or alternatively, to investigate learning capacities. Kahneman's psychophysical experiments during the 1960s on the visual system, on the other hand, were brief experiments in highly-controlled laboratories. The subjects had to distinguish between letters, digits or other visual stimuli, but a few observations per subject was generally considered sufficient. In Kahneman's experiments it was important that the value or "energy" was precisely the same for the experimenter and the subjects. But compared to Tversky's choice experiments, this meant that the experimental setting needed to be controlled as much as possible. Hence, Kahneman's visual experiments could only take place in laboratories which were specially built for this purpose.

In Kahneman and Tversky's collaborative experiments the small world requirement was abandoned. They no longer required that there be absolute certainty that the stimuli be understood in the same way by the experimenter and the subject. There did not need to be an actual draw from an urn or a toss of a dice in the presence of the individual, a hypothetical question about uncertainties sufficed. Repeated trials could be avoided on the grounds that most decisions were only taken once or twice, and not five hundred times in a row – although, admittedly, this point remained largely implicit. Furthermore, as the experiments could be done everywhere and all the time, no special laboratory controls were required. Answers to hypothetical questions obtained in the street from passers by were just as valid as the responses obtained from first-year students who had to participate in experiments to obtain their credits, or from subjects in controlled laboratories who were paid according to their performance. Finally, responding to hypothetical questions concerning decision problems produced enough evidence on which to base conclusions, and subsequently theory. Although Kahneman and Tversky considered experimental data based on real stakes better, answering hypothetical questions, in principle, sufficed.

There is not a direct link between the Heuristics and Biases program and the new, more relaxed standards of the experimental method. That is to say, Heuristics and Biases could, in theory, have been developed without the new means of conducting experiments. But the new standards of the experimental method facilitated Kahneman and Tversky's research in fundamental ways. It is safe to say that, without the method of hypothetical questions they could not have developed Heuristics and Biases. For instance, it would have been almost impossible to assemble eighty-four professional psychologists in a single laboratory in order to ask them what their opinions were on a statistical draw which was performed in front of them. Questions in which subjects were asked to choose hypothetically between a weekend in Paris or a week on a beach in Florida would have been impossible, if Savage's small-world requirement had still remained the standard. Thus, although there is not a direct link between Kahneman and Tversky's Heuristics and Biases theory and their experimental method, it is difficult to see how the theory could have been tested and developed without this new method.³⁰

4.3 Kahneman and Tversky versus Simon

It is tempting to view Kahneman and Tversky's heuristics as being similar to Simon's rules of thumb [cf. Heukelom (2007)] although this would be a mistake. In Simon's view, individuals use rules of thumb or heuristics to make decisions. An example of a heuristic could be when hearing the alarm clock in the morning, one gets up, takes a shower, and makes a cup of coffee. To Simon, the heuristic exists because the individual over time has learned that this is the best response to the stimulus of the alarm clock. In Kahneman and Tversky's approach, the function of heuristics was to simplify and reorganize the decision problem in such a way that it was manageable for a not very sophisticated decision maker. The heuristics determine how the new information of the stimulus is understood. The heuristics do not yield the decision, but reorganize the informational input in such a way that a decision making process is possible.

In the Linda problem, for instance, the individual intuitively believes it to be more likely that Linda is a bank teller and a feminist, as opposed to being just a bank

³⁰ I do not discuss here the question of whether different experimental methods yield different data and phenomena on which to construct theories. Obviously, someone such as Savage considered the experimental method crucial, whereas Kahneman and Tversky, as indicated above, considered it to be of much less relevance.

teller because he or she associates the information about Linda more with being a feminist than with a bank teller. In terms of the availability heuristic, 'feminist' is more available than 'bank teller' for the individual. Hence, the individual makes his or her intuitive decision on the basis of his or her understanding of the information presented. If it is given more thought, he or she may opt for the bank-teller-only option, particularly if he or she has just taken a course in logic. But the individual's intuitive initial response will always be the bank teller plus feminist option.

In addition, in Kahneman and Tversky's account the individual could not adjust his or her heuristics, as he or she could in Simon's approach. The Simon individual might replace coffee with orange juice when he or she learns that it is healthier. After a brief period in which extra effort is required to change the heuristic, the new heuristic will be to get up, take a shower, and drink a glass of orange juice. For Kahneman and Tversky, it appears to be the contrary, since the heuristics seem to be considered to be part of a given, unchanging human nature. Availability, representativeness, and so on, are seen as components belonging to the human information processing machinery that cannot be changed. They can be overridden by means of conscious, rational effort, but they always are determined by how the individual will behave when he or she is making decisions intuitively. To mathematical psychologists and behavioral decision researchers such as Coombs, Edwards, Luce and Tversky, a human, when acting intuitively, was acting as a statistician. In Kahneman and Tversky's research man was an intuitive statistician, an intuitive optimizer of utility, and an intuitive logician, although an imperfect one. The individual used the normative models, but only after the heuristics had reorganized the input. In many ways, this was a very different theory compared to Simon's theory of bounded rationality.

5. Prospect theory: Heuristics and Biases for economics

In 1979 Kahneman and Tversky published their famous article on "Prospect Theory: An Analysis of Decision under Risk" in *Econometrica*. The article marked a shift in emphasis away from probabilistic decision problems to an investigation of people's capacity to behave according to the normative theory of expected utility theory. It was the first attempt to produce a more complete descriptive theory of human decision making under uncertainty. Prospect theory has often been presented as being different

from Heuristics and Biases [e.g. Kahneman (2002)]. It is certainly true that prospect theory brought the different heuristics into one overarching framework, but the foundation still was the idea that human beings relied on a set of heuristics for their decision making and that the use of these heuristics sometimes leads to systematic deviations from the normatively correct decision. In this regard it is to be noted that it took Kahneman and Tversky some five years to get the article published in *Econometrica*, and that the last four of these five years were used to tweak a, for the most part, finished argument to fit an economic audience [Kahneman (2002)].

This continuity between Heuristics and Biases and prospect theory is illustrated by the remarks made in Tversky's letter to Krantz, April 10, 1975. In this letter Tversky, for the first time, devotes more than just one line to his scientific work with Kahneman, and he is clearly enthusiastic about the project. The letter illustrates that the basic argument of prospect theory had crystallized in the spring of 1975:

Danny and I are working primarily on decision making and we believe for the first time that we understand the basic principles governing choices between gambles. [...] The key elements in the theory we propose are: 1) an S-shaped utility function defined on differences from status quo rather than on total asset position and 2) uncertainty weights (not to be confused with subjective probability) by which the utilities are weighted. We are collecting empirical data which seem to provide very strong support for this model. [...] I will send you a draft of the paper in the very near future.

[Tversky's letter to Krantz, April 10, 1975]

In 1975, four years before prospect theory would be published, Kahneman and Tversky were in the middle of developing their Heuristics and Biases theory. Prospect theory, then, is best understood as an extended version of their Heuristics and Biases theory that focused on applications in economics. Its rhetoric was specifically designed to convince economists. Kahneman and Tversky's attempt to enter economics during this period is also illustrated by a workshop Tversky co-organized with Daniel McFadden, econometrician and 2000 co-winner of the Nobel memorial prize in economics. The workshop, held in October 1977 and entitled "Cognition,

Choice, and Economic Behavior,” was supported by the Mathematical Social Science Board and the National Science Foundation and was intended to bring together psychologists and economists interested in cognition and choice theory [McFadden and Tversky’s letter to Luce, June 20, 1977].

Kahneman and Tversky made the connection with their earlier work in the first few lines of the 1979 article, which set out the conception of expected utility theory as a normative theory which also makes descriptive claims:

Expected utility theory has dominated the analysis of decision making under risk. It has been generally accepted as a normative model of rational choice, and widely applied as a descriptive model of economic behavior. Thus it is assumed that all reasonable people would wish to obey the axioms of the theory and that most people actually do, most of the time.

[Kahneman and Tversky (1979), p.263].

In a clever way these opening sentences alluded to both the psychological and the economic framework. To psychologists these sentences restated the well-known normative-descriptive framework and signaled a contribution to an already established field of research. Positivist economists on the other hand might have raised their eyebrows at the injunction of the ‘normative,’ but they would certainly have agreed that reasonable people prefer to obey the axioms of expected utility theory and do so, or at least most of the time. Note, furthermore, that Kahneman and Tversky carefully avoided the term ‘rational,’ and used ‘reasonable’ instead. Evoking the term ‘rational’ might have suggested that this was an article in the line of critique of economics. The use of ‘rational’ would certainly have induced some economists to think that these two psychologists had the same research program as Simon, who had won the Nobel memorial prize the year before. From the start, prospect theory had been carefully constructed so as to be able to convince economists especially.

As in *Heuristics and Biases*, Kahneman and Tversky based their argument on a series of hypothetical questions they had presented to experimental subjects, in this case psychology students at Hebrew University. The problems the subjects were presented with were decision problems, involving different utilities and different

probabilities. Most of the questions were reformulations or variants of Maurice Allais' decision problems. After Allais had shown that even Savage himself had violated his own axioms, resulting in Savage making a distinction between a normative and an empirical domain, as set out in Chapter two, the "Allais paradox" in the 1960s and 1970s had become the iconic demonstration of a violation of expected utility [Jallais and Pradier (2005)].

One example of Kahneman and Tversky's use of an Allais-type approach is in the question where subjects were asked to state which of the following options they preferred.

A: (4,000, .80) or B: (3,000)

That is, they were asked whether they preferred 4,000 shekel with a probability of 0.8, or 3,000 shekel for certain.³¹ Most of the subjects in this case chose B. Assuming that the utility of the outcome equaled its monetary outcome this implied that they did not maximize expected utility. However, opting for the choice B could be the expected utility maximizing choice, if the decision maker was risk averse. Then subjects were asked which of the following two options they preferred.

C: (4,000, .20) or D: (3,000, .25)

In this case, most of the subjects chose C and hence maximized expected utility. This was problematic in combination with the first choice as it implied that subjects were sometimes risk averse, but on other occasions they maximized expected utility. Note that the second choice is equal to the first with probabilities divided by four. With such examples, Kahneman and Tversky illustrated that despite its normative status, expected utility theory as a descriptive theory was invalidated. In specific circumstances people systematically deviated from the norms of expected utility theory. A new descriptive, "alternative account of individual decision making under risk," was therefore required. The alternative account was christened "prospect theory" [Kahneman and Tversky, (1979), p.274].³²

³¹ At the time of the experiment, 4000 shekel was about one third of the modal monthly Israeli income.

³² Kahneman (2002) recalls that they deliberately looked for and chose a name that did not refer to any other theory or phenomenon in economics and psychology.

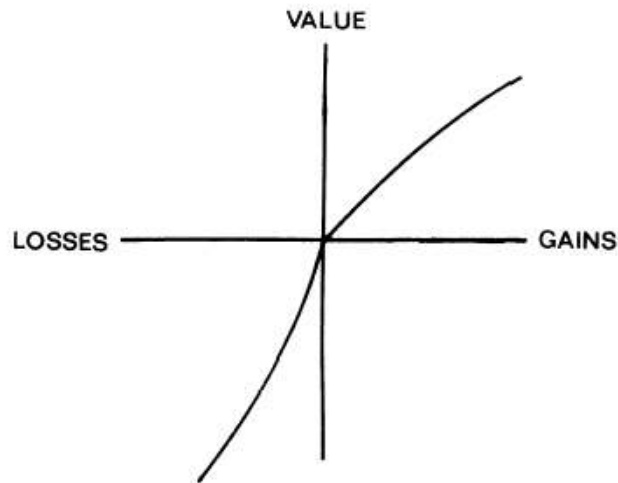
According to prospect theory, a human decision maker first employs a number of heuristics to make a decision problem manageable. This process is called the editing phase. Complicated decisions are broken down into different simpler decisions, different decisions are lumped together into one big decision, a benchmark is set with which the decision was compared, and so on. The purpose of this editing phase was to make the decision manageable. After this, the decision was evaluated in what was referred to as the evaluation phase. The evaluation phase had the same structure as the maximization of expected utility, but instead of the objective values of utility and probability, it used the individual's subjective perception of utility and probability. The subjective perception of utility was referred to as value (denoted v) and the subjective perception of probability was referred to as decision weight (denoted π). In expected utility theory, a subject who is faced with a choice between outcome x that occurs with probability p and outcome y that occurs with probability q derives utility according to the following function.

$$U(x, p; y, q) = p \cdot u(x) + q \cdot u(y) \quad (1)$$

In prospect theory, a subject that following the editing phase faces the exact same choice values this choice according to this function.

$$V(x, p; y, q) = \pi(p)v(x) + \pi(q)v(y) \quad (2)$$

Following the editing phase, value in prospect theory was a function of the outcomes; decision weight was a function of probability. The estimated relation, the relation based on the experimental results, between value and outcomes can be seen in the following graph.

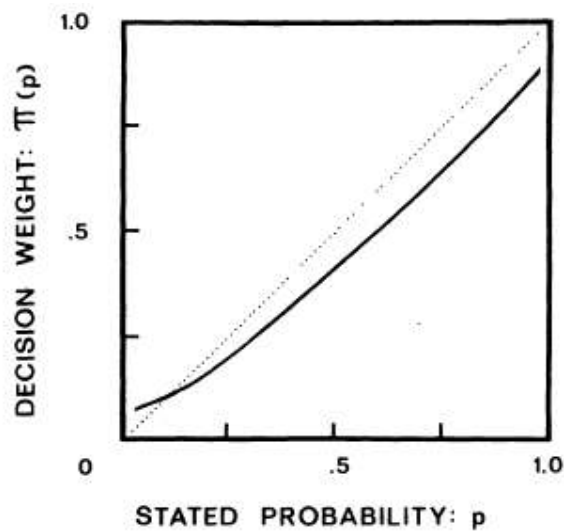


Kahneman and Tversky (1979), p.279

This relationship, in which the decrease in value from losses is larger than the increase in value from gains, held for each individual. It was, in other words, assumed that there was one functional form of the outcome-value relation that held for each human being.

At first sight it seems unnecessarily confusing to label the subjectively perceived outcome of a choice ‘value,’ instead of utility. But if we recall that in the second chapter it was set out that the Lewinian notion of ‘valence’ in psychology was used to denote the attractiveness or averseness of an object or choice option to the individual, we can see that by making ‘value’ instead of ‘utility’ the term that refers to the subjective attractiveness of a choice option, Kahneman and Tversky were able to design a framework that could be accepted by both psychologists and economists. To psychologists, the value framework matched the Lewinian valence framework; to economists it preserved the expected utility framework, while at the same time it allowed for the possibility that individual economic behavior deviates from the expected utility framework. In addition, the term ‘value’ was historically a central concept in economics that could be used very well to denote the pleasure an economic agent derives from choosing a specific option.

In prospect theory, individuals were believed to be similar in their subjective valuation of outcomes. Moreover, also regarding the relationship between probabilities and decision weights individuals were considered to be similar. Both the functional form and the numbers were equal for each individual. The estimated relation between decision weights and probability is depicted below.



The weighting function, Kahneman and Tversky (1979), p.283

In the figure, the dotted line represents the normative relation between the objective “stated probability” and the subjective “decision weight.” The dotted line connects the points in which the two are equal. This is the line one would find in the experimental measurement of the decision weights of the rational individual: for the rational individual perceived probabilities are the objective probabilities. The second, unbroken line shows the average of experimental measurements of reported decision weights of different objective probabilities by experimental subjects. The figure shows, that for small probabilities subjects over-perceive the objective probability, but for all probabilities above approximately 0.1, they under-perceive objective probabilities.

In Chapters two and three I discussed the connection between the theory of individual perception in psychology and the theory of scientific measurement. I recalled that they were two sides of the same coin in the days of Gustav Fechner, and showed that they continued to be two sides of the same coin in postwar mathematical psychology and behavioral decision research. However, now they were subject to the extra dimension of rationality. The value-utility and the decision weights-stated probability relations as presented by Kahneman and Tversky in their 1979 *Econometrica* article provides a clear example of this continued connection between the psychology of perception and the theory of measurement. First, the figures demonstrate how individuals perceive the objective stimuli of the choice, and thus

facilitate the theorizing of human decision making by demonstrating which information the human being uses to make his or her decision. But, second, the figures also show how the human measurement instrument deviates from its ideal functioning. It demonstrates how the human measurement instrument systematically deviates from the norm and provides a detailed relationship between the ideal measurement instrument and the actual measurement instrument. In other words, for prospect theory the psychological theories of human perception and decision making and theories of measurement are two sides of the same coin.

Kahneman and Tversky did not make this link explicit in their 1979 article in *Econometrica*, but a clear indication is that prospect theory is related to what is essentially psychophysics, although they did not use the term psychophysics as such. Kahneman and Tversky argued that “[a]n essential feature of the present theory is that carriers of value are changes in wealth or welfare, rather than final states,” and that “[t]his assumption is compatible with basic principles of perception and judgment.” This basic principle was that

[o]ur perceptual apparatus is attuned to the evaluation of changes or differences rather than to the evaluation of absolute magnitude. When we respond to attributes such as brightness, loudness, or temperature, the past and present context of experience defines an adaptation level, or reference point, and stimuli are perceived in relation to this reference point. Thus, an object at a given temperature may be experienced as hot or cold to touch depending on the temperature to which one has adapted.

[Kahneman and Tversky (1979), p.277]

Kahneman and Tversky extended this basic argument to other attributes: “The same principle applies to non-sensory attributes such as health, prestige, and wealth. The same level of wealth, for example, may imply abject poverty for one person and great riches for another – depending on their current assets” [Kahneman and Tversky (1979), p.277]. This principle of psychophysics, Kahneman and Tversky argued, was a fundamental aspect of economics.

The use of heuristics and the framework of psychophysics allowed Kahneman and Tversky to construct a theory in which individuals behaved rationally, and yet could often be observed as making irrational decisions. Man is rational, but because

human beings apply heuristics to reconstruct decision problems to manageable proportions, and because they have a specific perceptual system, their reasoned decisions may deviate from the normatively correct solution. Kahneman and Tversky had to cut the link between the normative and the descriptive theory in order to maintain the normative theory, while at the same time allowing for the conclusion that people systematically and persistently deviate from the norm. Human beings, who in Savage and Edwards' accounts were capable of rational reasoning, i.e. normal healthy adults, could no longer be expected to behave according to the normative rules. Therefore, also the arguments against the normative theory by normal rational people were potentially no longer valid arguments.

Prospect theory based its reasoning on mathematics. It took the mathematical principles of decision theory as the norm for behavior, and developed the mathematical measurement framework so that the experimental observations would fit. Deviations from the mathematical norms were understood as errors or mistakes, and they bore no implications for the norms. Because of the clear separation between the normative and the descriptive, it was now possible to construct a separate mathematical account of decision making in the descriptive domain, without making implications for the normative theory. In prospect theory, human beings were understood as having a biased perception of the relevant input of uncertainties and utilities, just as they had a biased perception of sensory inputs such as temperature and weight. Because of these imperfections, Kahneman and Tversky argued, their behavior would often deviate from the optimal norms of normative decision theory.

6. Economics and psychology

In economics, prospect theory has been understood as an attempt to apply a psychological theory to an economic question. It has been viewed as a theory taken from another discipline that has traveled across the scientific border in order to plead its arguments in a neighboring discipline. In other words, prospect theory has been understood by economists in terms of their own understanding of the relationship between economics and psychology. This has been set out Chapter one and will be further discussed in Chapters five, six, and seven. However, working as psychologists, Kahneman and Tversky understood economics, insofar as it concerned individual decision making, as a subfield of psychology. Moreover, following Edwards they regarded behavioral decision research as a psychological research program that

derived much of its inspiration from economics. Hence, Kahneman and Tversky regarded prospect theory, not as the application of a psychological theory to a question in the neighboring field of economics, but as an attempt to unify psychology and economics, and in particular as an attempt to unify behavioral decision research and economics. To acquire an understanding of prospect theory's reception in economics, it is therefore important to understand first how Kahneman and Tversky themselves understood prospect theory's relevance for economics.

6.1 Normative versus descriptive

In prospect theory Kahneman and Tversky employed their normative-descriptive distinction, and assumed that economists would employ the very same distinction. This partly reflected the standard understanding of economics in behavioral decision research and psychology. It should not be forgotten that Edwards understood economics' theories of individual human behavior as being normative theories. A valid conclusion would therefore be that Kahneman and Tversky were not familiar with the economists' use of positive and normative, and thus assumed them to mean the same as descriptive and normative in behavioral decision research, therefore employing the terms as they had been accustomed to.

But Kahneman and Tversky's use of normative and descriptive can also be seen as a very clever way of trying to convince economists of the relevance of prospect theory for economics. Note, that in their earlier *Heuristics and Biases*, their changing use of normative and descriptive played an important role. In prospect theory, Kahneman and Tversky did not tell the economists that their theory was complete nonsense or useless. Instead, they claimed to understand economics as using one theory to cover both the normative and the descriptive realm. For the normative part they fully agreed with economists, which fitted in neatly with practice in behavioral decision research. But, Kahneman and Tversky argued, economists had been mistaken in using that same theory in the descriptive domain.

Kahneman and Tversky's approach differed in a subtle but fundamental way from Simon's, the other main psychological critic of economics. Just as Kahneman and Tversky, Simon understood economics to have both normative and descriptive ambitions, but unlike Kahneman and Tversky, he considered economics to have embarked on the wrong track entirely. According to Simon, economics failed to distinguish correctly between descriptive and normative, had an inappropriate theory

in both domains, had an absurdly restrictive notion of rationality, and was much too narrowly focused on the mathematical advancement of its theories.³³

Kahneman and Tversky were much less hostile. In fact, they were in favor of the current practice in economics – after all, behavioral decision research and measurement theory were considered to be at least partly based on economics – and they only meant to suggest that a few adjustments be made to improve it. Contrary to Simon, Kahneman and Tversky argued that there was nothing wrong with the economists' theory of expected utility maximization. It was only that this was the normative theory, and not an accurate description of actually observed human behavior. Economists did not need to abandon the theory of expected utility maximization, but instead they should seek a proper descriptive counterpart to this normative theory. Prospect theory was then suggested as serving as such a descriptive theory.

6.2 Prospect theory as unification of economics and psychology

Prospect theory was aimed at economists. Unlike Heuristics and Biases, which was aimed at psychologists, prospect theory aimed to make an argument in economics. It was, however, a specific type of argument. Kahneman and Tversky were not attempting to travel across the psychology-economics border, to become economists and to make a contribution to economics. What they intended to do, was rather to shift the economics-psychology border in such a way that their work and economics would become part of the same science. Subsequently, they could then argue that their paper had proved existing theories wrong, and had provided a viable alternative. For their argument it did not really matter whether one understood the move as shifting the border so that parts of economics became part of psychology and behavioral decision research, or as shifting the border so that behavioral decision research became part of economics. The message would remain the same, namely, that behavioral decision researchers and economists were all part of the same scientific program, and that although prospect theory showed that many economists had been partially mistaken, the problem had been solved. With prospect theory, Kahneman and Tversky made a claim of unification; they implicitly argued that behavioral decision research and economics were really one and the same.

³³ e.g. Simon (1956, 1959, 1987). For overviews and discussions of Simon's position regarding economics see Sent (2005), Augier and March (2004), and Heukelom (2007).

In the discussion of prospect theory above a number of instantiations of this claim to a unification of economics and psychology have been mentioned. To begin with, the first few lines of the 1979 prospect theory article drew a distinction between normative and descriptive that was acceptable to both psychologists and economists, and thus bridged the gap between psychologists' normative-descriptive framework and economists' positive-normative distinction. Psychologists could read the remark about expected utility as having both normative and descriptive claims as referring to the standard framework in behavioral decision research. Economists could read the remark as saving the utility maximizing framework upon which all their theories were built, while at the same time opening the door to behavior that deviates from utility maximizing. Second, Kahneman and Tversky argued that the main problem of economics was that it did not include psychophysics in its account of individual behavior, in effect implying that economists should study individual behavior the way psychologists did. This meant that economists had mistakenly believed that economics and psychology were different disciplines, where in fact the two used the same behavioral foundations. Third, the distinction Kahneman and Tversky made between the objective 'utility' and the subjective 'value' of a choice appealed to both psychologists and economists. For psychologists it meant that they could understand prospect theory in the common Lewinian valence framework whereas to economists it was acceptable because Kahneman and Tversky had stayed within the bounds of the different ways in which these terms were used in economics.

The reason that prospect theory became so successful was that it had succeeded in combining and using conceptual frameworks from behavioral decision research and economics in such a way that scientists involved in decision making in economics found it useful. Economists could account for empirical anomalies without having to sacrifice their theories. Critics of economics could refer to prospect theory when arguing that economics was descriptively wrong, and behavioral decision researchers felt justified in their belief that economics was directly involved in, if not part of behavioral decision research.

7. Assessing Kahneman and Tversky

Throughout their collaborative career Kahneman and Tversky met with a range of opposing arguments.³⁴ As I will show in Chapter six, economists who initially enthusiastically adopted the Kahneman-Tversky framework in the 1980s, in the 1990s and 2000s gradually departed from this framework. Yet, in spite of this, their work has been tremendously influential over the years. Moreover, part of the opposition they have met with can probably be explained by their success and influence itself.

With their conceptual framework, Kahneman and Tversky offered a rationale for science as the ultimate foundation of rational decision making and of scientists as the ultimate experts of rationality. This has been arguably the most important part of the theoretical framework that Kahneman and Tversky exported to economics, and has given rise to behavioral economic paternalism, to be dealt with in Chapter six. For Kahneman and Tversky, as well as for the behavioral economics that grew out of their research, scientists were the ultimate experts on rationality, and thus ultimately decided what a rational decision is and is not.

In nineteenth and twentieth century psychophysics and experimental psychology the scientist was the ultimate expert regarding the objective value of the stimulus because it was the scientist who determined the value of the stimulus. Experimental psychologists wanted to know how an individual perceived the different stages in the brightness of a light bulb, and in these experiments experimental psychologists naturally knew what the objective values of the different stages of brightness were because they themselves had set up the experiment. Savage and

³⁴ The most often repeated claim has been that Kahneman and Tversky believed people to be irrational. The argument was that if human beings can send people to the moon and return them safely, they cannot be that irrational. One of the most remarkable exponents of this view has been Edwards, who wondered how it is possible that people are so poor assessing uncertainties, as they were in Kahneman and Tversky's theory, and yet, at the same time, could be so skilled in driving their cars. Edwards has never elaborated upon his reservations with respect to the work of Tversky and Kahneman, perhaps from fear of losing the image of coherence of his program, yet it is no secret that he disagreed with their work [e.g. Phillips and Von Winterfeldt (2006), Krantz – interview (2008), Dawes – interview (2008)]. The 1980s and 1990s have given rise to a whole surge of criticism regarding Kahneman and Tversky's approach to human decision behavior. The most prominent philosophical critique was provided by Cohen (1981). The most extensive criticisms from within experimental psychology came from Gigerenzer [see e.g. Gigerenzer (1991, 1993, 1996), Gigerenzer and Murray (1987), Sedlmeier and Gigerenzer (1997, 2000), Hertwig and Gigerenzer (1999)]. These are summarized and discussed in Heukelom (2005)]. Gigerenzer has been the only critic to whom Kahneman and Tversky have explicitly responded [Kahneman and Tversky (1996)]. Other critics include Lopes (1991) and Cosmides and Tooby (1996).

Edwards applied this experimental psychological program to decision making, but limited the superior knowledge of the scientist. Savage emphasized that all normal healthy adults could evaluate the axioms of rational decision making, and assumed that everyone would agree with his axioms after some careful thought. Edwards adopted Savage's framework, and he equally assumed that human beings in principle make their decisions in accordance with the axioms of rationality.

Following the repeated experimental evidence of violations of rationality in Tversky's experiments in the 1960s, combined with the outcome of Kahneman's psychophysical research that human beings are fundamentally flawed decision makers, Kahneman and Tversky in the 1970s developed a theory that assumed that human beings often and systematically deviate from what is rational and normative. Heuristics and Biases and prospect theory detached the normative from the descriptive, and constructed a theory for understanding what happened in the descriptive domain. One consequence of detaching the normative from the descriptive was that normal healthy adults were no longer qualified judges of the axioms of the normative theory in rational decision making. The only person still qualified to judge whether a specific decision was rational or not was the scientist who possessed a thorough training in logic, statistics and decision theory.

Although prospect theory was about decision making, the link with measurement theory can still be detected. The link shows up in the utility-value curve and the stated probability-decision weight curve very clearly. These curves describe how the individual perceives the objectively given utility and probability of choice, and they thus provide the basis for understanding decisions in the descriptive domain. However, by describing how individuals perceive objective stimuli in the form of utilities and probabilities, it also provides an account for the deviations found in the human measurement instrument, and thus provides a correcting factor that ensures the possibility of using the human being as a measurement instrument.

In different ways, the joint work of Kahneman and Tversky constituted a break from Tversky's earlier work on decision making and measurement theory of the 1960s, although it was a relative break. In his joint work with Kahneman, Tversky took a different approach to human decision making and, as a consequence, took a different approach to measurement. No longer did he maintain that human beings on average are rational decision makers, and no longer did he assume that the human

measurement instrument on average is unbiased. But decision making and measurement were still two sides of the same coin. What had changed was that Tversky now understood human beings to deviate systematically and predictably from what is normatively correct.

5. What to conclude from psychological experiments?

How Smith and Thaler incorporated behavioral deviations in economics

1. From psychology to economics

In the late 1970s economists who followed Vernon Smith's (1927-) experimental work in economics, corroborated the experimental results of behavioral decision research and concluded that rational choice theory had not been successful in describing individual economic choice behavior [Grether and Plott (1979, 1982)]. Shortly thereafter, financial economist Richard Thaler (1945-) used these experimental findings to infer very different implications for economic theorizing [Thaler (1980, 1993)]. Thaler introduced the work of Daniel Kahneman and Amos Tversky, and argued it provided the solution to the empirical anomalies encountered in financial economics. Thaler's efforts contributed to the development of behavioral finance in the 1980s, which led to the rise of behavioral economics in the 1990s. The rise of behavioral finance and behavioral economics forced Smith to clarify his position on individual economic behavior. In the 2000s, this produced a collaboration between Smith and Gerd Gigerenzer (1947-) and it was through Gigerenzer that Smith would align experimental economics with the work of Simon.

This chapter thus deals with a transition period in the 1980s in which new methods and insights from behavioral decision research altered both experimental and mainstream economics. The second section of this chapter retraces the initial corroboration of the experimental results of the psychologists by experimental economists in the late 1970s and early 1980s. The third section shows how Thaler and a number of other financial economists accepted the experimental corroboration but inferred a different conclusion. The fourth section sets out Thaler's continued efforts to contribute to developing the field of behavioral finance in the 1980s. Section five discusses Smith's response to the rise of behavioral finance. Section six serves as a conclusion.

2 Corroboration and incorporation of psychology's behavioral deviations in Smith's experimental economics

In the 1960s and 1970s Smith gradually came to the conclusion that economics needed to be altered [e.g. Lee (2004), Smith (1962, 1965, 1967, 1974)]. His experiments formed an important basis for this. Smith stressed that time was

necessary for the market to reach an equilibrium and argued that experiments should be used to investigate which factors in the real world determine to which equilibrium the market drives the economy to over time. According to Smith, economics was too theoretical and failed to look seriously at actual behavior in the real-world economy. Smith complained that the standard references, Paul Samuelson's *Foundations of Economic Analysis* (1947) and Roger Allen's *Mathematical Analysis for Economists* (1938), only discussed "the purely formal properties of the theory" [Smith (1959), p.65], and were of little direct use when applied to real-world problems. He complained that these authors talked about the "inputs" of the production function without giving them any interpretation. When one did so one immediately was forced to make a distinction between the different kinds of inputs, Smith argued, and as a result one ended up with quite different mathematical results. Smith insisted repeatedly that, as opposed to the standard theory, his position had implications "in a very real economic sense" [Smith (1959), p.67].

Smith's experimental results and his growing dissatisfaction with modern economics led him midway through the 1970s to what is probably his strongest denouncement of this framework. "I believe that the microeconomic theory of the pre-1960's is a dead end," Smith wrote, and immediately added an alternative: "The new microtheory will, and should, deal with economic foundations of organization and institution, and this will require us to have an economics of information and a more sophisticated treatment of the technology of transacting" [Smith (1974), p.321]. However, Smith did not imply that pre-1960s microeconomics should be put aside, but argued instead for a serious revision and extension of the theoretical framework. Smith had his reservations with respect to the economics of Allen and particularly the economics of Samuelson and he developed an experimental method that many of his fellow economists were not yet ready to accept.

Smith's link to behavioral decision researchers and other psychologists was close. Smith's important early collaborator, Sidney Siegel (1916 – 1961), was a psychologist [e.g. Innocenti (2008)]. Experimental research in economics more generally arose from a brief postwar period of cooperation between operations researchers, computer scientists, mathematicians, economists and psychologists [Dimand (2005), Weintraub (1992), Lee (2004)]. The link was, for instance, close enough for Smith to participate in discussions on Leonard Savage's and others' decision theory. In "Measuring Nonmonetary Utilities in Uncertain Choices: The

Ellsberg Urn” (1969) Smith took a position in the ensuing debate on the violations of Savage’s normative theory as presented by Maurice Allais, Daniel Ellsberg and others. Smith’s response to the Ellsberg argument is important because it shows how Smith attempted to strike a balance between Savage’s theory and its opponents, an attempt in which he tried to bridge the opposing theoretical and experimental sides. “I stand with those, like Savage, Raiffa, and Schlaifer, who say they would not want to violate the axioms consciously,” Smith started his argument. Yet, he was unwilling to go all the way with Savage: “However, having stated this I am not prepared to assert that he who seriously and consciously violates the axioms, and in my judgment knows what he is doing, is thereby simply making a ‘mistake,’ and should be given a little more conditioning and ‘education’” [Smith (1969), pp.324-325].

Smith took a position that in a crucial way differed from the position taken by decision theorists and behavioral decision researchers. Smith did not want to violate Savage’s axioms, but he did not accept the conclusion that people who violate the axioms were making mistakes either. People may have very good reasons for deviating from the axioms, Smith argued. For instance, this could be because they take into account what other people, such as friends and colleagues, think of their decisions. Smith did not accept the conclusion that deviations from the axioms are to be understood as mistakes implying the need for education. For Smith, deviations from Savage’s axioms, even when they were systematic, were not problematic because over time the market would correct those mistakes. Whereas decision theorists, such as Savage and behavioral decision researchers such as Ward Edwards and Amos Tversky believed that systematic violations immediately raised questions about the normative theory, for Smith this link was much less direct. For Savage, Edwards and Tversky, a decision was either normatively correct (rational) or normatively false (irrational). To Smith, rationality was a matter of content and degree. People might have reasons for initially deviating from the norms, and in a market context the institution of the market would ensure that in due time they would adjust their behavior towards the rational behavior. Smith was of the opinion that “even if [Savage’s] axioms are to be regarded as basically a normative theory, the theory can also do valuable service in helping us to understand actual behavior” [Smith (1969), pp.324-325]. The normative theory shows where and when people deviate from the norm, and in that sense guides the description of observed decision

behavior. But it also serves as a description of human behavior in market equilibrium, and thus helps us to understand how decision behavior adjusts over time in markets.

Another way in which Smith's experimental work differed from that conducted by behavioral decision researchers is that his decision makers were not individual subjects but economic units. In one of his theoretical papers on investment and production planning, Smith, for instance, started as follows: "We imagine individual decision making units, which we call 'firms'" [Smith (1960), p.198]. The individual decision making units of Smith's theories and experiments were not individual human beings, as in the decision theory and the experiments of the psychologists, but they were individual economic units such as firms, consumers, and producers. Smith was not interested in the individual as an individual, but was interested in the individual in its role as a particular economic decision making unit.

As matter of fact, Smith was not even interested in the individual as an economic decision making unit, but instead in how the market institution influenced the unit's behavior over time. In "Experimental Studies of Discrimination Versus Competition in Sealed-Bid Auction Markets" [Smith (1967)] Smith stated that the "primary purpose" of his experiments was "to study individual bidding behavior and price determination under two alternative forms of market organization: (1) price discrimination, [...] and (2) pure competition" [Smith (1967), p.56]. It was not the individual unit's behavior that should be investigated, but the market environment that affected its behavior. In mathematical psychology and behavioral decision research the individual functioned as a measurement instrument for (average) individual psychological characteristics, as set out in Chapters two and three. In Smith's experiments, the individual functioned as a measurement instrument for characteristics of the market mechanism.

Because Smith's experiments are historically and methodologically connected to the experiments of the psychologists of behavioral decision research, and because Smith at different points actively engaged in discussions in decision theory, it might appear that experimental economics, decision theory, and behavioral decision research developed in tandem in the 1960s and 1970s. But Smith's experimental economics differed in at least two crucial ways from the experiments conducted by the psychologists. First, Smith did not investigate the individual human being, but instead he investigated economic decision making units. Smith was only interested in individual human beings in their role as an economic decision making unit. Second,

Smith did not assume the static point of view in which it is believed that the individual that deviates from the normative theory has made a mistake. In contrast, Smith was interested in how decision behavior changes over time, and in the environments that induce these changes. Smith took a stance that he at least once labeled “a crude macrobiological approach,” in which the system, when not exogenously altered, tends towards a “stable equilibrium” [Smith (1968), p.410].

But Smith did not entirely dismiss the experimental results of the psychologists either. The stream of experimental results obtained by behavioral decision research in the 1970s which showed that individuals violate rational choice theory, required experimental economists assume a stance. David Grether (1939-) and Charles Plott (1938-), who had joined Smith’s experimental economics program in the 1970s, decided to subject the experimental results of the psychologists to a test. The reason that they could do so was that all the material and expertise were already available. The rise of experimental economics had produced an environment in which the results of the psychologists could be tested without requiring economists to learn new methods or techniques. Checking as many possible explanations as they could think of for the results obtained by the psychologists, Grether and Plott (1979, 1982) sought to falsify the findings of the psychologists. Moreover, they set out to test the experimental findings based on the presumably much more rigorous standards of (experimental) economics. The rise of the experimental method in economics had made experimental economists confident they could beat psychologists at their own game, or at least critically assess their work by using their own experimental method.

Grether and Plott (1979, 1982) focused on the alleged phenomena of “preference reversals,” the phenomenon that occurs when individuals change their preferences regarding the same choice when it has been differently formulated; and “intransitivity,” the related phenomenon showing that actual individual preferences are not always transitive.³⁵ Grether and Plott (1979) were very suspicious of the empirical evidence produced by the psychologists, and aware that economists and

³⁵ Their references to behavioral decision research consisted of only a few articles [e.g. Slovic and Lichtenstein (1971, 1973), Tversky (1969, 1971)]. Paul Slovic and Sarah Lichtenstein were, however, quick to remark in the *American Economic Review* that “there is a substantial body of research on preference reversals within the psychological literature that is being neglected here. Moreover, reversals should be seen not as an isolated phenomenon, but as one of a broad class of findings that demonstrates violations of preference models” [Slovic and Lichtenstein (1983), p.597]. From that moment on, investigations of empirical falsifications of rational choice theory in experimental economics increased.

psychologists did not always use a rational choice theory for the same purpose. They emphasized that “[t]here is little doubt that psychologists have uncovered a systematic and interesting aspect of human choice behavior” [Grether and Plott (1979), p.624], but wondered whether 1) the phenomenon also held in more typical economic situations, and 2) whether it could be explained by means of economic theory.

The main worry of Grether and Plott was that the experimental results were mere artifacts produced by the experimental setup of the psychologists. They produced thirteen (!) methodological and theoretical economic explanations for the falsifications: 1) no real money was used and incentives may therefore have been misspecified; 2) different incomes of the subjects may have influenced some experiments; 3) in most of the psychological experiments indifference between two options was not possible; 4) perhaps subjects did not give their true selling or bidding price but acted strategically; 5) subjectively perceived probabilities from the lotteries used may not be equal to actual objective probabilities; 6) perhaps subjects chose lexicographically, as in Tversky’s elimination-by-aspects theory [Tversky (1971)], which would account for a moderate form of preference reversals; 7) perhaps the magnitudes of the choices were too close, leading to apparent intransitivity, as in Tversky (1969); 8) the cost of decision making could be too high compared to the expected pay-off, leading subjects to not make an effort; 9) perhaps the choices subjects faced contained too much information for the subject to process within the time available; 10) subjects could have been confused or might have misunderstood the experiment; 11) perhaps the phenomena reported occurred only in a few subjects; 12) the subjects were relatively unsophisticated psychological undergraduates, whereas more sophisticated subjects might make more rational choices; and 13) the experimenters were psychologists, leading subjects to speculate about the true purpose of the experiments, and hence perhaps to change their behavior.

This last explanation for the findings particularly illustrates that Grether and Plott went to great lengths to show that the findings of the psychologists had been mere artifacts. Grether and Plott’s message was that every possible explanation for the psychologists’ findings needed to be controlled for, even the argument that the results should not be taken seriously for the sole reason that the experimenters had been psychologists. Grether and Plott set up two experiments in which they controlled for all thirteen possible explanations. They specified incentives, they made the experiments very simple and they made certain that all subjects understood the

choices they could make. Furthermore, they used undergraduates as well as graduates, making it clear that they were economists and not psychologists, and they took the two possible explanations of Tversky into consideration. But, much to their surprise, they obtained results that were similar to those of the psychologists. Consequently, they remained “as perplexed as the reader who has just been introduced to the problem” [Grether and Plott (1979), p.624].

The first Grether and Plott article was published in the same year as Kahneman and Tversky’s prospect theory, and in Grether and Plott (1982) they recognized prospect theory as a prominent example of a rational choice theory which adjusted a number of assumptions in order to account for the empirical findings. But Grether and Plott (1982) stressed that prospect theory could not account for their experimental results; “We need to emphasize that the phenomenon causes problems for preference theory in general, and not for just the expected utility theory. Prospect theory as a special type of preference theory cannot account for the results” [Grether and Plott (1982), p.575].

The conclusions Grether and Plott derived from these results are important because they set the standard for experimental economists’ responses to these and similar findings for the following quarter of a century. According to Grether and Plott (1979, 1982), the experimental results pointed to an inconsistency between actual behavior and rational choice theory that was “deeper than the mere lack of transitivity or even stochastic transitivity.” The empirical results suggested “that *no optimization principles of any sort* lie behind even the simplest of human choices” [Grether and Plott (1979), p.623, emphasis added]. Grether and Plott did not believe that the empirical results could be addressed by making a relatively minor adjustment of rational choice theory, but drew the radical conclusion that utility maximization and rational choice should be completely abandoned as a description of and explanation for the decision making behavior of individuals.

However, and this was equally crucial for the experimental economic approach, Grether and Plott did not imply that utility maximization and rational choice as a description of market behavior were invalidated. With respect to market behavior the experimental results only showed that the economic subjects, who in the final market equilibrium behave according to rational choice and utility maximization, initially behave according to a to-be-developed theory that is completely unlike utility maximization and rational choice. Because the disciplining, *rationalizing* institution

of the market operates between individual behavior and market behavior, a falsification of individual rational optimization did not falsify rational choice as a description of equilibrium market behavior. Quite the contrary, the experimental results only emphasized the role of the market as the mechanism that rationalizes individual behavior. Smith, Grether and Plott assumed that the market drives the behavior of the economy to a rational, utility maximizing equilibrium over time. Therefore, the fact that initially individual behavior systematically deviated from rational utility maximization only showed how important the market mechanism was in driving individuals to rational behavior in market equilibrium.

3. Thaler's financial economic anomalies and the creation of behavioral finance

In postwar neoclassical economics, the market had largely become an empty concept, mainly due to Samuelson's work. In Samuelson's framework [Samuelson (1947)], it was assumed that all the decision makers in the economy always maximize utility. Samuelson needed to make this assumption for his operationalist approach. If individuals were assumed not to always maximize utility it could not be supposed that their observed behavior was on individual demand and supply curves and it would then be impossible to operationalize demand and supply curves through the operation of measuring individual choices. But assuming that individual subjects in the economy always maximize utility implied that the economy was always in equilibrium [Weintraub (1991)]. In contrast to the view later developed by Smith and in experimental economics, time was not an element in the Samuelsonian neoclassical economic world. Samuelson only considered static equilibria. There was a direct link between the behavior of individuals and the market; the market was nothing more than the sum of all the individual behaviors. The adjustment of individual behavior to market equilibrium did not come into play.

In the 1960s and 1970s, a new field in economics appeared that used the Samuelsonian neoclassical theory as a theoretical foundation for its empirical investigation of stock market behavior. Based on research conducted during the 1950s and 1960s by Franco Modigliani (1918 – 2003), Merton Miller (1923 – 2000), and Harry Markowitz (1927-), financial economics, as the new field came to be called, gradually appeared as an accepted genuine sub-branch of neoclassical economics in the second half of the 1960s and 1970s [Jovanovic (2008), Poitras and Jovanovic

(2007), MacKenzie (2006)].³⁶ The empirical study of stock markets was linked to neoclassical economics through what came to be referred to as the efficient market hypothesis. The efficient market hypothesis specified the theoretical position of neoclassical economics in the case of the stock market, and any market for that matter [Jovanovic (2008)]. “A market in which prices always ‘fully reflect’ available information is called ‘efficient’” [Fama (1970), p.383].

The central question for financial economists was whether stock markets indeed are efficient, as theory predicted, or inefficient, for which an explanation then would have to be found. In the second half of the 1960s and 1970s two opposing views developed. At MIT, Paul Cootner (1930 – 1978), Hendrik Houthakker (1924 – 2008) and others developed and defended the idea that the stock market was not efficient. “[P]rice changes are not purely random but follow certain longer run trends,” Houthakker argued [Houthakker, quoted in Jovanovic (2008), p.228], and Cootner more bluntly stated that “[t]he stock market is not a random walk.” [Cootner, quoted in Jovanovic (2008), p.225]. The Chicago Graduate School of Business held and fiercely defended the opposite view, that the stock market is efficient and that the stock prices over time will appear to be a random walk [Jovanovic (2008)]. In Chicago, Eugene Fama (1939-), a student of Miller, arose as the main protagonist, defending the efficient market as an empirical and theoretical fact [e.g. Fama (1970)]. Inspired indirectly by then Chicago mathematician Savage, Fama distinguished between “sophisticated traders” who were experienced enough to determine the intrinsic value of securities and act accordingly, and “other participants” who did not (yet) possess this skill and who produced the random noise around the intrinsic value. Sophisticated traders ensured that the market prices remained or returned quickly to the underlying value of the stocks [Jovanovic (2008), Fama (1970)].

On the basis of the empirical corroboration of the psychological results by Grether and Plott, financial economists began to look seriously at the results of the psychologists. One illustrative example is Arrow (1982), which discussed a number of phenomena in the (stock) market that contradicted the “rationality hypothesis,” such as individuals’ unwillingness to accept government subsidized insurance below its actuarial value and observed irrationality in financial markets. Arrow suggested “that these failures of the rationality hypothesis are in fact compatible with some of the

³⁶ See e.g. Markowitz (1952, 1959, 1965), Modigliani and Miller (1958), Miller and Modigliani (1963) and the collected papers in Modigliani, Abel, and Johnson (1980).

specific observations of cognitive psychologists” [Arrow (1982), p.5]. Arrow, in other words, drew a direct line from observations in the laboratories of the psychologists to contradictions observed in the market. Experimental results from psychology, that in themselves had nothing to do with the economy or with markets, were linked to economics and used as an explanation for the unsolved financial economic puzzles. In one sentence Arrow linked two very different phenomena:

an important class of intertemporal markets shows systematic deviations from individual rational behavior and [...] these deviations are consonant with evidence from very different sources collected by psychologists. [Arrow (1982), p.8]

Systematic deviations from rational behavior by individuals in the laboratory could be an explanation for observed market deviations only when one understood the relation between individual and market behavior to be direct, as was the case in Samuelsonian neoclassical economics.

However, deviations from the theory of efficient markets had been discussed before. For instance, market deviations could be explained as resulting from market imperfections such as transaction costs and limited information. The deviations could furthermore be explained as short-run phenomena that would quickly disappear through arbitrage. But Arrow considered these explanations to be insufficient and argued that the results of the psychologists and experimental economists implied that market imperfections should be understood as genuine phenomena, and not as having resulted from temporary distortions of the market. Thus, Arrow recalled that “[a]ny argument seeking to establish the presence of irrational economic behavior always meets a standard counterargument: if most agents are irrational, then a rational individual can make a lot of money; eventually, therefore, the rational individual will take over all the wealth” [Arrow (1982), p.7]. But, Arrow argued, arbitrage and related arguments could easily be countered: “(1) Not all arbitrage possibilities exist. [...] (2) More important, if everyone else is “irrational,” it by no means follows that one can make money by being rational, at least in the short run” [Arrow (1982), p.7].

Attempts to incorporate the corroborated findings of the psychologists in financial economics appear scattered through the literature from 1980 onwards. But

Thaler was the first economist to draw economic implications from behavioral decision research findings explicitly. This central focus of his work made him a great promoter of Kahneman and Tversky's work in economics. Extensive references to the work of Kahneman and Tversky occurred in almost every publication by Thaler. Thaler was an economist from the Chicago Research School of Business and a colleague of Fama. In the 1980s he worked predominantly in financial economics, advancing the experimental results and the theoretical approach of Kahneman and Tversky as an explanation for the observed falsifications of the efficient market hypothesis, and thus disagreeing with the prevalent financial economic view in Chicago. In the 1980s the exploration of systematic deviations from the efficient market hypothesis in financial markets by Thaler and others became known as behavioral finance.

Thaler's first behavioral finance paper, "Toward a Positive Theory of Consumer Choice," appeared in 1980 in the *Journal of Economic Behavior and Organization*. By 1991 Thaler had collected enough material to publish a book, entitled *Quasi Rational Economics* and consisting of sixteen of his papers that tested the traditional neoclassical economic models and offered alternatives. In 1993 Thaler edited another book entitled *Advances in Behavioral Finance* for the Russell Sage Foundation (RSF), consisting mainly of papers from the latter half of the 1980s, which was followed by a second volume in 2005, with the same title. Kahneman and Tversky were behavioral finance's theoretical founding fathers, but Thaler was its earliest and strongest advocate.

Specifically, Thaler built on two lines of Kahneman and Tversky's research. Thaler systematically connected Kahneman and Tversky's biases of rational choice in experiments to the anomalies of rational choice theory found in financial economics, and he made this connection the cornerstone of a new research program. He collected phenomena that were anomalous in financial economics and were compatible with the biases found by Kahneman and Tversky. Sometimes explanations were offered on the basis of prospect theory or by means of some other theory. Usually, however, these violations of standard economic theory were presented without any explanation to account for them, and he simply stressed what they implied, that neoclassical theory had been violated.

For the *Journal of Economic Perspectives*, Thaler published two series of “anomalies” columns that had the sole purpose of proclaiming that economics had serious problems. The first series contained fourteen anomalies articles and appeared between 1987 and 1991.³⁷ The second series contained four publications and appeared between 1995 and 2001. The first anomaly article in 1987 documented “the January effect.” When the market for stocks is in efficient equilibrium, in the neoclassical world the average monthly return should be equal for each month. There is no reason to expect that stocks would perform better just because it happens to be a certain month. However, this was exactly what was observed in the case of January. Especially for smaller firms stock returns were substantially higher in January compared with other months. How could this January effect be possible given the theory of efficient markets? The answer was that it was not possible and that one needed a theory such as Kahneman and Tversky’s prospect theory to account for the findings.

Loewenstein and Thaler (1989) showed that many similar anomalies existed in and outside the economy that have to do with intertemporal choice. For example, people prefer to pay too much tax in advance and to receive some back when the year is over instead of the reverse, even when the first option is subject to costs in terms of lost interest. Schoolteachers who can choose between being paid in nine months (September-June) or in twelve (September-August), choose the second option although from an economic perspective the first is more rational. But Loewenstein and Thaler also cited the dermatologist who lamented that her patients were unwilling to avoid the sun when she told them about the risks of sun cancer, but who were quick to stay out of the sun when she told them about the risk of getting “large pores and blackheads.” This example, Loewenstein and Thaler argued, was also a violation of economic theory because it showed myopia in patients they should not have if they acted rationally. The implicit reasoning was that economic theory could be applied to every aspect of our lives and that therefore also violations of economic theory could be drawn from every corner of life. The recurring message of the anomalies articles was that there are serious problems with economic theory which cannot be easily dismissed, and which need to be taken seriously.

³⁷ The anomalies of the first series have been collected in *The Winners Curse* (1992).

In his anomalies column Thaler cited examples from finance that were clearly economic. The structure of the anomalies was often similar to the biases produced by Kahneman and Tversky. One anomaly that Thaler frequently investigated and that became one of the principal anomalies of behavioral finance was the “endowment effect.” The endowment effect was an application of the framing effect of Kahneman and Tversky that showed that individuals’ preferences are subject to an initial framing process. In other words, individuals’ preferences depend on the quantity of the means they are endowed with. The experiment is as follows. Divide a group of subjects randomly into two sub-groups and give one of the two sub-groups a standard coffee mug. Subsequently, ask the sub-group with the mug what price they would minimally want to sell the mug for. Also ask subjects of the sub-group without mugs what price they would maximally want to pay for the mug. Typically, the willingness to accept (WTA) is about twice the willingness to pay (WTP). Apparently, people reframe their preferences after receiving the mug. In economics, this endowment effect could serve as an explanation for the often observed fallacy of taking into account sunk costs [see e.g. Thaler (1980, 1987), Tversky and Kahneman (1981)]. The endowment effect further falsified the Coase theorem, which says that in order to attain the efficient market allocation, the initial endowment of the goods should be irrelevant. The Coase theorem depends on the assumption that for every individual WTA equals WTP, so that trading will continue until the goods are in the hands of those with the highest WTP. But given the demonstrated systematic difference between WTA and WTP, the Coase theorem no longer held true [Kahneman, Knetsch, and Thaler (1990)].

A defining characteristic of Thaler’s behavioral finance was that it adopted Kahneman and Tversky’s understanding of normative and descriptive. In addition, Thaler accepted their understanding of the positive realm of economics as covering both the normative and the descriptive domain. This understanding diverged from the postwar economic understanding in which a value-free positive domain was contrasted with the application of positive economic theories for specific value-laden goals of the policy maker in the normative realm. As a result, Thaler’s introduction of Kahneman and Tversky’s meaning of normative and positive could not but lead to confusion. Thaler equated Kahneman and Tversky’s descriptive domain with the

economists' positive and used normative both in Kahneman and Tversky's meaning and in the economists' meaning. Thaler (1980) expressed it as follows:

Economists rarely draw the distinction between normative models of consumer choice and descriptive or positive models. Although the theory is normatively based (it describes what rational consumers *should* do), economists argue that it also serves well as a descriptive theory (it predicts what consumers in fact do). This paper argues that exclusive reliance on the normative theory leads economists to make systematic, predictable errors in describing or forecasting consumer choices. [Thaler (1980), p.39]

The conceptual re-organization of economics that Thaler took over from Kahneman and Tversky played an important role in behavioral economics in the 1990s and 2000s and it would determine how behavioral economists started to think about policy advice in the 2000s. This will be discussed in more detail in Chapter six. In line with Kahneman and Tversky, Thaler argued that further theoretical advancement of the normative theory was perfectly fine, but that because economists had ignored the fact that real-world behavior of individuals does not agree with this theory so long, they should now also pay more attention to building a descriptive theory of economic behavior.

Thus, Thaler not only accepted the empirical evidence presented by Kahneman and Tversky, but he also accepted their accompanying methodological distinction. Essentially, Thaler accepted Kahneman and Tversky's attempt to recreate economics in the image of behavioral decision research. As set out in Chapter four, prospect theory made a unificatory claim. It claimed that behavioral decision research and economics were part of the same program, and that the approach of behavioral decision research was better than that of the economists. Therefore, the economists should adopt prospect theory and its methodological distinctions. Thaler accepted this reasoning entirely. He provided Kahneman and Tversky's approach with more economic content, but he left the theoretical structure intact. From the early 1980s to the early 1990s, Thaler's promotion of Kahneman and Tversky's prospect theory

acted as the main catalyst for establishing an economic program based on the work of Kahneman and Tversky.

Another catalyst for developing behavioral finance in the 1980s was the support of the Alfred Sloan Foundation (ASF), the Russell Sage Foundation (RSF) and the National Bureau of Economic Research (NBER). The best way to conceive of the role of the NBER, the ASF and the RSF is in contributing to the attachment of a small group of researchers to a large and influential research program. Their financial resources explain how a few financial economists interested in research from a particular branch of psychology could develop a more or less coherent research program built largely around the work of two psychologists. From 1986 onwards, the ASF and later the RSF were consistent sponsors of behavioral economic research. In the mid-1990s the RSF set up a series of books in behavioral economics, set up a “Behavioral Economic Roundtable” that regularly brought and still brings behavioral economists together, and organized and still organizes a series of workshops in collaboration with the NBER. The financial support from ASF, RSF and NBER is a relevant characteristic of the rise of behavioral finance and behavioral economics, but is not further discussed in this dissertation.

4. Distinguishing experimental economics from the rising star of behavioral finance

The difference between Thaler’s behavioral finance and Smith’s experimental economics is that behavioral finance investigates individual behavior and that experimental economics investigates markets.³⁸ The growing number of behavioral financial publications in the 1980s and the influence of Kahneman and Tversky’s work more generally pressed Smith to distinguish his experimental economics more clearly from these psychologists and their economic off-spring. In 1989, ten years after the first Grether and Plott article, Smith asked:

³⁸ From a different starting point, Ana Santos’ dissertation *The Social Epistemology of Experimental Economics* (2006) arrives at a taxonomy that is similar to, yet differs somewhat from the history described here. Within what she broadly labels “experimental economics” she distinguishes three types of experiments: “market experiments” such as those of Smith, and “non-market experiments,” the latter including “individual behaviour and decision-making experiments” such as those of Allais (1953) and “game theory experiments,” such as Kalish, Milnor, Nash, and Nering (1954) and Schelling (1957).

How do we close the [...] gap, between the psychology of choice and agents' economic behavior in experimental exchange markets? [...] I think we economists need to accept these replicable empirical results [of behavioral decision research] as providing meaningful measures of *how people think about economic questions*. For their part, psychologists need to accept the dominating message in experimental research on the performance of a wide variety of bidding, auctioning and customer (posted price) markets: markets quite often “work” in the sense that over time they converge to the predictions of the economists' paradigm. [Smith (1989), p.165, emphasis in original]

The conclusions drawn from the experimental results of Grether and Plott (1979, 1982) have been held by Smith, Grether, Plott and other experimental economists from the late 1970s until today. Over the years experimental economists have struggled over how to formulate their approach and how to distinguish their ideas from Thaler's behavioral finance and more broadly from those of the behavioral economists. Part of the difficulty was (and still is) that experimental economists and Kahneman and Tversky's prospect theory are seemingly very close. Experimental economists agree that the psychologists' experimental findings indeed disprove rational choice of individual decision behavior, which easily led to the conclusion that they also agreed with the theoretical implications that were drawn by behavioral decision researchers and behavioral finance economists.

Another difficulty was that experimental economists conducted the same kind of experiments as the behavioral decision researchers and behavioral finance economists, but with a different purpose. Behavioralists conducted experiments with individual human subjects to investigate the decision making characteristics of the individual. Experimental economists conducted experiments with individual human subjects to investigate the market. The two sides conducted the same experiments, but with a different question in mind. Chapters two to four showed that experimental psychology used human subjects not because it was interested in any particular individual, but as a measurement instrument for measuring the characteristics of *the* individual. Experimental economics went a step further. It was not interested in the particular individuals in experiments, nor in the individual or his or her characteristics in general. It needed the individuals to experiment on a phenomenon that was altogether different from the individual subjects of the experiment. Like the biologist

who investigates a virus through its effect on laboratory mice, so experimental economics investigated the market but needed individual subjects as it were, to investigate the market. Nevertheless, experimental economists were easily understood as investigating human behavior, as will be set out in more detail in the following chapter. Frequently found statements in experimental economics of the sort “[i]n laboratory market experiments, we test the theory’s assumptions about agent behavior” [Smith (1989), p.154] could understandably be misunderstood as statements about the psychology of human beings. Moreover, as a result of this subtle distinction, experimental economists were pressed to distinguish themselves more clearly from Thaler’s behavioral finance and later from behavioral economics.

The use of time was crucial in Smith’s experimental economics. In experimental economics the market required time to drive the economy to equilibrium. Because of their use of time experimental economists could maintain that individual behavior initially deviated from the norms of rational choice theory and utility theory. At the same time, they could also maintain that the emerging market equilibrium was in line with rational choice and utility theory. Smith’s use of time distinguished him sharply from Samuelsonian neoclassical economics and behavioral economics and this was the main reason why he had difficulties explaining experimental economics’ position to behavioral finance economists. *“People have their own homegrown beliefs about how markets work, or should work,”* Smith carefully explained, and *“questionnaire responses reflect these beliefs, which are often couched in terms of ‘fairness’ criteria.”* As a consequence, *“[people’s] initial behavior in a market may reflect these beliefs.”* However, when these individuals operate in a market over time their behavior *“adapts to the incentive properties of markets”* [Smith (1989), p.166, emphasis in the original].

Smith went on to conclude that in economics there were “two experimental research programs,” both of which, he added, required considerable development. First there was the “economist’s maximizing paradigm,” which “often performs well in predicting the equilibrium reached over time in experimental markets.” However, the economist’s maximizing paradigm “is not generally able to account for short run dynamic behavior, such as the contract price paths from initial states to final steady states.” Second, there was “the psychologist’s ‘reference frame’ descriptive paradigm,” which is Kahneman and Tversky’s prospect theory and the behavioral finance that emerged from it. This psychological program did well “in explaining

subjects' introspective responses, and their short-run or initial decision behavior, but it provides no predictive theory of reference frame adjustment over time." Smith was quick to point out that a well-known paper from the psychological program agreed with this analysis. "In fact, the statement (Kahneman, Knetsch and Thaler, 1986, p.731) 'that they (people) adapt their views of fairness to the norms of actual behavior' can be interpreted as a description of what is observed in experimental markets" [Smith (1989), p.166].

Another way in which Smith tried to distinguish experimental economics more clearly from behavioral decision research, behavioral finance and behavioral economics was by abandoning the label "experimental game" as a description of his experiments. In the 1950s, 1960s and early 1970s game theory, the application of rational choice theory to situations of human interaction, had been an important source of inspiration for Smith's experiments [Lee (2004), Weintraub (1992), Dimand (2005)]. But game theory as a description and explanation of the interaction of rationally acting self-interested individuals started from a description of individuals as optimizers of utility. The experimental results showed that this had been a wrong assumption. Therefore, game theory became inappropriate for experimental economists as a description of individual behavior. The fact that Smith during the 1970s explicitly discarded his use of the term "experimental games" to describe his experiments seems an unsolved puzzle [Lee (2004)]. But in the light of the corroborations produced by Grether and Plott (1979, 1982) Smith's reasons can be illuminated. Smith considered game theory no longer a good description of individual behavior. Game theory still describes and explains the behavior of the individuals in the eventual market equilibrium, but cannot explain individuals' behavior when they are first presented with an economic decision. It can neither explain the process of adjustment to equilibrium. As a result, the term "experimental game" became inappropriate as a description of experiments that investigated the adjustment behavior of the individual agents in a market setting. The experiments were still considered a game in the sense that they mimicked the crucial aspects of the market, but they were no longer a game in the sense of describing fully rational interacting individuals.

Pressed to distinguish experimental economics more clearly from behavioral finance, Smith was led to cooperate with psychologist Gigerenzer in the 2000s. In 2001 Smith contributed a chapter to Gigerenzer and Reinhard Selten's edited volume

Bounded Rationality, The Adaptive Toolbox [McCabe and Smith (2001)]. In 2008 Smith published a monograph entitled *Rationality in Economics, Constructivist and Ecological Forms* in which he drew an explicit link between his own and Gigerenzer's work. In the *Handbook of Experimental Economic Results, volume 1* (2008), edited by Plott and Smith, Gigerenzer participates by making no fewer than six contributions, and the third, forthcoming volume in Gigerenzer's Adaptive Behavior and Cognition (ABC)'s research group on bounded rationality in its title emphasizes the link with Smith: *Ecological Rationality: Intelligence in the Real World*. Finally, in 2008 Smith agreed to become co-director of Gigerenzer's ABC group at the Max Planck Institute for Human Development in Berlin [email Gigerenzer to author, July 12, 2008].

5. Economics, behavioral decision research, and Kahneman and Tversky

In the 1960s and 1970s behavioral decision research became relevant to economics, as set out in Chapters three and four. Behavioral decision researchers produced experimental results that seemed to falsify rational utility maximizing behavior, the basis of all modern economics. The role of Kahneman and Tversky has been emphasized: Kahneman and Tversky's prospect theory (1979) was a conscious attempt to influence economics and to alter economists' reasoning.

The first to pick up on behavioral decision research's experimental results were experimental economists such as Grether, Plott and Smith. Grether and Plott (1979, 1982) corroborated the experimental findings of psychology and drew the conclusion that rational choice as a description of individual human behavior should be entirely abandoned. However, experimental economists at the same time concluded that rational choice as a description of efficient markets in equilibrium could be maintained and that the experimental results only emphasized the rationalizing forces of the market. Furthermore, experimental economics did not accept behavioral decision research's alternative accounts, and explicitly denounced the most visible theory among them, Kahneman and Tversky's prospect theory.

An unexpected result of experimental economists' corroboration of behavioral decision research's experimental results was that it paved the way for behavioral decision researchers to enter financial economics. Thaler understood the psychological findings to show the irrationality of individual choices, and drew a direct link from the irrationality of individual choices to irrational features of the

behavior of markets. He immediately recognized Kahneman and Tversky's research, and especially Kahneman and Tversky (1979), as an important new and improved theory of individual decision behavior. The different responses of Smith's experimental economics and Thaler's financial economics to the experimental results and to prospect theory's alternative can be explained in terms of the different notion of the market in experimental economics and financial economics. To Smith the market was a rationalizing mechanism that requires time to drive the economy towards equilibrium. For financial economists such as Thaler, time was not an element of the market.

This different response to the findings of behavioral decision research was a reflection of a more fundamental difference between experimental economists and behavioral decision researchers such as Kahneman and Tversky. Behavioral decision researchers located the explanation for the deviations from rational behavior in the nature of human beings. For Kahneman and Tversky the reason that human beings often deviate in their behavior from what is rationally optimal is because they are made that way. As set out in Chapters two, three, and four, Kahneman and Tversky stood in an experimental psychology tradition in which the fixed characteristics of *the* individual were investigated. Kahneman and Tversky investigated the human being in the way that the physicist investigates the atom: on the assumption that there is one universal way in which the individual/atom can respond to a particular stimulus. This was a fundamentally different understanding from that of experimental economists who started from the assumption that individual behavior is constantly subject to change and who investigated how the market causes individual behavior to change. Experimental economists investigated behavior like biologists in that they assumed that individual behavior constantly adapts to an external selection force.

This also explains why prospect theory was so quickly adopted by many economists, whereas Simon had largely failed to influence economists. Kahneman and Tversky, neoclassical economists like Samuelson, and financial economists like Thaler, all conceived of individual behavior as a stable phenomenon that, like the physicist's atom, could be isolated to investigate its fixed and universal properties. Simon and Smith on the other hand were social scientists who understood individual behavior as adaptive and hence as unfixed. In their view, therefore, it did not make sense to isolate the individual decision maker from its environment. The universal properties Kahneman and Tversky, Thaler, and Samuelson were looking for simply

did not exist in this view of economic behavior. The two sides had a different understanding of individual behavior and a different understanding of collective decision making in markets, administrative organizations and institutions. Simon and Smith clashed with mainstream neoclassical economists on a very basic level while Kahneman and Tversky agreed with financial economists on the fundamentals.

Financial economists had an approach not unlike behavioral decision research and therefore were not as fundamentally opposed to behavioral decision research as were experimental economists. But not being opposed is not the same as favoring. At least two further reasons explain why Kahneman and Tversky specifically were successful in influencing many mainstream economists. First, Heuristics and Biases and prospect theory explained the violations of the efficient market hypothesis as argued by financial economists in the 1960s and 1970s. Yet, as rational choice theory was a cornerstone of financial economics, abandoning it completely, as Simon for instance had proposed, would be too radical a step. Kahneman and Tversky's prospect theory was successful because it offered rational choice theory an honorable way out. Kahneman and Tversky's message to economists was that there was nothing wrong with their theory of rational choice, but that economists should recognize that their label of positive actually covered a normative and a descriptive domain. Kahneman and Tversky explained that economists had been using rational choice theory for both the normative and the descriptive domain, whereas it should only be used for the normative realm. Kahneman and Tversky thus offered financial economists one straightforward way out of their problems.

Second, the alternative Kahneman and Tversky offered in the form of prospect theory was close to rational choice theory and in fact essentially was rational choice theory as seen through psychophysical spectacles. Being close to the traditional rational choice theory had the advantage that it was the same framework for individual behavior that financial economists had been using and were familiar with. The conceptual step from traditional rational choice theory to prospect theory was a small one. Furthermore, prospect theory did not require serious alterations or that the theories and models economists used be abandoned. The introduction of a few extra parameters sufficed to allow financial economics proceed as before.

An unfortunate element that has blurred the understanding of this episode in the history of economics and psychology is that both the alternative proposed by Simon in the 1950s-1970s and the group of Kahneman and Tversky-inspired

economists that arose in the 1980s and 1990s have been labeled behavioral economics, a label first used by George Katona in 1946 [Juster (2004)]. This has led to the misunderstanding that the behavioral finance and behavioral economics of Thaler was the continuation or a reappearance of the earlier project of Simon. But this is not true. Simon's behavioral economics and Thaler's behavioral economics were two very different projects, and the idea that they share the same label, and thus must be somehow related is a red herring.³⁹

The late 1970s and early 1980s constitute a transition period in the recent history of economics. Smith and experimental economists gained prominence and used their new method to critically test the experimental results of behavioral decision research. They found that behavioral decision researchers had been right in important respects and concluded that rational choice theory was not successful as a complete description of human behavior. Financial economist Thaler, by contrast, drew a very different conclusion and reconceptualised financial economics on the basis of Kahneman and Tversky's behavioral decision research. This unified financial economics with a specific branch of psychology, since in behavioral finance, economics became part of behavioral decision research. We will see in Chapter six that, from the late 1990s onwards, behavioral economists would again distinguish themselves from psychology and from behavioral decision research in particular. But in the 1980s, Kahneman and Tversky's claim that behavioral decision research and economics were essentially about the same thing had been wholly accepted and applied in Thaler's behavioral finance.

³⁹ A classification such as 'old' and 'new' behavioral economics of Sent (2004) only provides easy ammunition to those eager to show that the two behavioral economics' are related. Thus, it is common in contemporary behavioral economics to refer to Sent's distinctions to quickly show that contemporary behavioral economics incorporates Simon's behavioral economics [e.g. Angner and Loewenstein (forthcoming a,b)].

6. Building and defining behavioral economics

1. Who are we?

George Loewenstein, a prominent behavioral economist, recalls that

In 1994, when Thaler, Camerer, Rabin, Prelec and I spent the year at the Center for Advanced Study in the Behavioral Sciences, we had a meeting to make a kind of final decision about what to call what we were doing. Remarkably, at that time, the name behavioral economics was not yet well established. I actually advocated ‘psychological economics,’ and Thaler was strong on behavioral economics. I'm kind of glad that he prevailed; I think it's a better, catchier, label, although it creates confusion due to association with Behaviorism. [email Loewenstein to author, June 16, 2008]

It was no accident that Richard Thaler gave the new field its name. Thaler's behavioral finance had developed in the 1980s as a sub-field in financial economics, but towards the late 1980s and early 1990s Thaler realized that behavioral finance could be broadened to include other violations of neoclassical economics. One example is the anomaly paper of Thaler and Loewenstein [Loewenstein and Thaler (1989)] discussed in the previous chapter. It started with the violation of rational intertemporal choice in financial markets, but subsequently also applied the theory to subjects' behavior in the dermatologist's office. As a consequence of this broadening of behavioral finance, more economists and non-economists began to join Thaler's research program. The central argument was that neoclassical economics was descriptively wrong in every domain to which it was and could be applied. This development culminated in the creation of the field of behavioral economics in 1994, as Loewenstein recalls in the above quote.

During the 1990s and 2000s, Thaler and other behavioral economists expanded behavioral economics from a small research program focused on violations of the neoclassical theory in financial economics to a dominant research program that looked for inspiration beyond behavioral decision research to a range of scientific disciplines and methods. To date, they have investigated violations in every aspect of life and from the early 2000s onwards they have developed their own stance on policy advice under the rubric of libertarian, light, or asymmetric paternalism. As a result of

this expansion, behavioral economists came to realize how behavioral economics relates to neighboring fields. Thus, in the late 1990s and 2000s, behavioral economics started to distinguish itself more explicitly from neighboring fields, such as experimental economics and psychology. All of these developments helped to expand behavioral economics into a broad and stable economic research program in which the influence of Daniel Kahneman and Amos Tversky relatively declined. The two psychologists remained the iconic founding fathers, but the omnipresent influence they had achieved in Thaler's behavioral finance in the 1980s no longer existed. Kahneman and Tversky's normative – descriptive distinction remained the methodological basis of behavioral economics, but the labels were changed into full rationality versus less-than-fully, quasi, or, eventually, bounded rationality. This change of terminology made behavioral economists redefine the history and subject matter of their field, no longer claiming that only Kahneman and Tversky were their founding fathers, but Herbert Simon as well.

This chapter presents the story of how behavioral economics was built and redefined in the 1990s and 2000s. Of course, the literature in behavioral economics is so vast that it is impossible to do justice to the wide variety of research within the scope of one chapter. I shall, therefore, proceed by discussing the examples which can serve to illustrate the move to expand and redefine the field in crucial ways. Research into intertemporal choice and emerging preferences challenged standard assumptions in neoclassical theory about the time consistency of choice, and the given character of preferences. Using Kahneman and Tversky's distinction between a normative benchmark and descriptive deviations, this research moved in some instances even so far as to challenge the normative benchmark of rational choice itself. As we shall see, this last move was for some behavioral economists a 'bridge too far.' The research into preference formation shows how a split emerged between those economists who did not accept the normative benchmark and who moved into the direction of research such as that of psychologist Gerd Gigerenzer (under the label of "ecological rationality"), and those economists who kept adhering to the normative-descriptive distinction, albeit in modified form. Both strands claimed, as later will be seen, Simon as their founding father, but for different reasons. We will see that behavioral economists used the terminology of bounded rationality to explore the rapidly growing field of neuroeconomics. Here, we see the independent influence that using new tools of research had on economists' practices as well.

Behavioral economists gradually built their program into a stable and well-defined mainstream economic program. In this process, the main question was how to construct the descriptive theory of human decision behavior. To answer that question behavioral economists explored a range of different scientific disciplines and methods. At the same time, however, they remained faithful and always came back to the normative – descriptive framework that had originally been introduced by Kahneman and Tversky. This conceptual core determined how behavioral economists understood the economic world, it determined the welfare implications they drew, and finally, it determined how they pulled back when their explorations diverged too far from this conceptual core. The new terminology of rationality lay the foundation for discussing behavioral economists' new paternalistic stance on economic policy advice which developed from the early 2000s. The building of behavioral economics from the mid 1990s onwards led behavioral economists to distinguish themselves more clearly from psychology and experimental economics. Redefining their own field in the process, behavioral economics is now one of the most thriving and innovative branches of economics as a discipline.

2. Building behavioral economics

2.1 Intertemporal choice and the dual system approach

Behavioral economic research on intertemporal choice started in the early 1990s and culminated in the behavioral economic research based on the two-systems approach. The intertemporal choice and two-systems approach literature is illustrative for a number of reasons. First, it was chronologically the first major theme behavioral economists focused on and it continues to this day to be an important topic in behavioral economics. Second, it illustrates how behavioral economists incorporated Kahneman and Tversky's work, and in particular their normative – descriptive distinction into economics. During this process behavioral economists sometimes would venture so far as to also question the normative benchmark, but in the end they always pulled back to the conceptual foundation as laid out by Kahneman and Tversky. Finally, it illustrates how behavioral economists developed a theoretical framework that could be applied to any economic problem and was compatible with important developments in neuroscience and the cognitive sciences in general.

The two most prominent behavioral economists who have worked on intertemporal choice have been George Loewenstein and David Laibson. Loewenstein

finished his PhD at Yale in 1985 and published his first article in 1987. From his first publication onwards he has been a strong proponent of more psychology in economics, but initially he was hardly influenced by the work of Kahneman and Tversky. Instead, an important theoretical influence came from the work of Jon Elster, with whom he wrote several articles and edited a book for the Russell Sage Foundation called *Choice over Time* (1992).⁴⁰ Loewenstein has published a number of articles on the history of psychological and economic explanations of intertemporal choice and utility, revealing an extensive knowledge of the history of economic thought [e.g. Loewenstein (1992), Elster and Loewenstein (1992), Frederick, Loewenstein, and O'Donoghue (2002), Angner and Loewenstein (forthcoming a,b)]. Laibson finished his PhD at MIT in 1994 and started his academic career at Harvard University that same year. At MIT and Harvard he has focused on violations of the traditional economic idea of exponential discounting. His articles are a mixture of experimentally corroborating this phenomenon, building mathematical economic models that account for the observed systematic deviations, and investigating the psychological and neurobiological substrates of the observed behavior.

For Loewenstein the problem of the well-known exponential discounting utility (DU) model was not just that individuals discount hyperbolically, but it goes even further. For instance, individuals can be shown to sometimes use a negative discount rate [Loewenstein and Prelec (1991)]. If individuals prefer an increasing real-wage over a constant real-wage, even when the present value of the latter is higher than the former, they effectively employ a negative discount rate. Perhaps even more challenging for received economic theory was that individuals' intertemporal choices could be shown to be fundamentally inconsistent [e.g. Prelec and Loewenstein (1997)].⁴¹ People who prefer *A now* over *B now* also prefer *A in one month* over *B in two months*. However, at the same time they also prefer *B in one month* and *A in two months* over *A in one month* and *B in two months*. In other words, when faced with an intertemporal choice individuals like to save the best for last, which is in fundamental disagreement with economic theory. Another, by now famous descriptive falsification of the DU model, is the research on New York City cab drivers who judge their income "one day at a time" [Camerer, Babcock, Loewenstein, and Thaler (1997)].

⁴⁰ See for a discussion on Elster's work on decision making e.g. Davis (2003).

⁴¹ Here, Loewenstein referred to Friedman and Savage (1948, 1952) who face an analogous problem in explaining both gambling and insurance behavior.

The DU model fails not only descriptively, but also normatively, Loewenstein argued. For instance, there does not seem to be a good reason to suppose that somebody who is indifferent towards oranges and apples today should also be indifferent towards 1) apples today, oranges tomorrow, and apples the day after and 2) apples three days in a row. Loewenstein argued that there is little normative and descriptive reason for holding on to the DU model, despite its aesthetic merits of mathematical simplicity and consistency. This conclusion produced tension in Loewenstein's work. On the one hand, he was an early proponent of Thaler's behavioral finance, and a founding member of behavioral economics. But on the other hand, he concluded that economic theory might also be problematic as a normative benchmark.

Not surprisingly, then, Loewenstein was ambivalent about how to proceed. A number of publications show fundamental problems with the DU model both descriptively and normatively [e.g. Loewenstein (1992, 1999), Loewenstein and Prelec (1991), Prelec and Loewenstein (1997)]. But on different occasions Loewenstein also tried to extend the DU model as a descriptive theory which could fulfill its role while maintaining the normative benchmark. For instance, he built a mathematical model that can accommodate observed behavior while maintaining the normative benchmark as a limiting case [see e.g. Loewenstein and Prelec (1992)]. The discount factor is generalized to $1/(1+at)$, where a can be exogenously given or determined by another function. Loewenstein also turned his attention to neuroscience as a possible means to a solution [see e.g. McClure, Laibson, Loewenstein, and Cohen (2004)].

Loewenstein's explorations led him to doubt the normative – descriptive distinction of behavioral economics in the 1990s, but after a while he pulled back from contesting the normative benchmark and made his work compatible with Kahneman and Tversky's approach. In the 2000s, Loewenstein acknowledged that his work had been stimulated “by the existence of a strong normative benchmark, expected utility theory,” that behavioral theory in economics and psychology had advanced, and that “both psychologists and economists have made important theoretical and empirical contributions” [Loewenstein, Weber, Hsee, and Welch (2001), p.367]. Moreover, he added that in this area the “convergence in the theoretical perspectives of psychologists and economists [...] has been greater than for

any other topic of mutual interest in the two disciplines [Loewenstein, Weber, Hsee, and Welch (2001), p.367]. This use of Kahneman and Tversky's normative and descriptive in behavioral economics will be discussed in more detail below in Section 2.3.

A similar tension can be observed in the work of Laibson. In Laibson (1997), "Golden Eggs and Hyperbolic Discounting," he built a mathematical model of agents with hyperbolic discount functions that could explain a myriad of dynamically inconsistent individual preferences observed in experiments. "Golden Eggs" referred to the traditional rational economic individual decision model. In Laibson's model, the individual was faced with an "imperfect commitment technology," such as a retirement plan, which required that it be initiated one period before it started to work. Together with the hyperbolic discount function this model "predicted" that individuals' consumption would closely track the progress of their income, but that with the "imperfect commitment technology" individuals were capable of correcting their hyperbolic discount functions by committing themselves in advance to their desired savings behavior. Because the imperfect commitment technology required individuals to commit themselves in advance, the far-sighted, rational planner effectively constrained the temptation to be immediately gratified once the money actually arrived.

Ipso facto the model predicted that with "financial innovation" savings rates would go down because commitment technology no longer needed to be started up a period in advance. According to Laibson, this provided an explanation for the ongoing decline in U.S. saving rates. "Financial innovation" should be interpreted broadly here. It not only comprises new saving plans at banks, but also changes in "social commitment devices" such as marriage, work and friendship. The idea, which went back to Laibson's dissertation, was that a decrease in the strength of the structure and/or duration of long-term social commitments increased the probability of acting according to the short term hyperbolic discount function.

Furthermore, Laibson also showed that the dynamically inconsistent preferences resulting from hyperbolic discounting might lead to a welfare reduction following financial innovation. Financial innovations allowed individuals' short term hyperbolic discount functions to override their long term rational discount functions. Under certain conditions, the result may, from a rational, long-term perspective, be a reduction in welfare.

Laibson's 'Golden Eggs' article is a typical 1990s contribution to behavioral economics. It made productive use of Kahneman and Tversky's distinction between the normative and the descriptive by reinterpreting this in terms of a "far-sighted" and a "myopic" planner. Whereas in the 1980s Thaler would have used accessible language and would have referred to psychology to explain and solve the problem, Laibson presented the problem in a formalistic economic language and did not make any reference to psychology. Behind this emphasis on economics, however, Laibson was fully committed to the normative – descriptive distinction of Kahneman and Tversky. Golden eggs and hyperbolic discounting are the normative and descriptive rephrased in economic language. However, Laibson also added something to this framework, namely the idea that in the economy individuals might have the possibility of controlling their deviating behavior by means of commitment technologies. This idea of commitment technologies extended the Kahneman and Tversky framework. It suggested that individuals might be both deviating from the normative benchmark, while at the same time helped to explain how they deviate from the norm. Further, it suggested that it is not so much the scientists that need to find ways to correct individuals' deviating behavior, but the individuals themselves.

Harris and Laibson (2001), "Dynamic Choices of Hyperbolic Consumers," elaborated further on the idea of hyperbolic discounting. It tried to link the short term hyperbolic discounting with the long term (rational) exponential discounting and showed how individuals act who try to prevent their own future overconsumption. The paper started with the traditional discounting function for individuals, and replaced the constant discount factor δ with an "effective discount factor." This effective discount factor consisted of the sum of two components, the "long-run discount factor δ " and the "short-run discount factor $\beta\delta$," where hyperbolic discounting implied $\beta < 1$. The traditional discount factor was hence explicitly decomposed into a long-run, exponential component and a short-run, hyperbolic component. The assumption was that individuals, faced with "stochastic income" and a "borrowing constraint," anticipate their future inclination to hyperbolically discount (and thus to overconsume), and that they want to act against it. Hyperbolic discounting was thus explained as resulting from a strategic game with future selves.

Since $\beta < 1$, the effective discount factor is negatively related to the future marginal propensity to consume (MPC). To gain intuition for this effect, consider a consumer at time 0 who is thinking about saving a marginal dollar for the future. We assume that this consumer acts strategically in an *intrapersonal* game where the players are temporally situated “selves.” The consumer at time zero – ‘self 0’ – expects future selves to overconsume relative to the consumption rate that self 0 prefers those future selves to implement. Hence, on the equilibrium path, self 0 values marginal saving more than marginal consumption at any future time period. From self 0’s perspective therefore, it matters how a marginal unit of wealth at time period 1 will be divided between savings and consumption by self 1. Self 1’s MPC determines this division. Since self 0 values marginal saving more than marginal consumption at time period 1, self 0 values the future less the higher the expected MPC at time period 1. [Harris and Laibson (2001), p.936, emphasis in the original]

In equilibrium, self 0 would reduce his or her savings rate (and thus his or her future income) to the point where his or her preference for self 1’s savings rate would be equal to self 1’s actual savings rate. In other words, because individuals knew they would discount hyperbolically in the future, they would also discount hyperbolically now. In equilibrium the two selves maximize the combination of their preferences. The effective discount rate was a function of a “time preference” (the difference between preferences of self 0 and self 1) and an anticipation of future MPC.

Harris and Laibson (2001) also demonstrates how Laibson retreated from the idea that individuals can influence their own behavior through commitment. Instead, the behavior displayed was merely the result of conflicting selves, none of whom could be controlled by the other. Like Loewenstein, Laibson struggled between exploring new directions and remaining committed to a distinction between the norm and its imperfect realization.

The research on intertemporal choice illustrates how behavioral economists extended Thaler’s behavioral finance. First a piece of standard neoclassical economics was examined; in this case the assumption of exponential discounting. Subsequently, it was shown that this piece of the neoclassical theory failed descriptively. In the next step, the piece of neoclassical economics as descriptive theory was adjusted to be

compatible again with the empirical facts, the same approach as used in Kahneman and Tversky's prospect theory.

The behavioral economists' way of dealing with intertemporal choice described human behavior as the outcome of two systems or processes striving for dominance. Different labels appear in the behavioral economic literature for these two systems: reasoning vs. intuition [e.g. Kahneman (2003)], rationality vs. emotion [e.g. Shefrin and Thaler (1988), van Winden (2007), Ben-Shakar, Bornstein, Hopfensitz, and van Winden (2007)] and cognitive vs. affective [e.g. Camerer, Loewenstein and Prelec (2005)] are the most prominent. Understanding human behavior as the outcome of conflict between different motives has a long and rich history, going back to the philosophy of Plato and Aristotle and to Homer's Ulysses tying himself to the mast so he could hear the Sirens sing [Davis (2003), pp.63-80]. Behavioral economists' re-creation of neoclassical economics' understanding of individual behavior in terms of two souls inhabiting one body is therefore a recent development in a long history.

Some behavioral economists have linked this dual system solution to research in neuroscience and neurobiology, thus contributing to the creation of a new sub-field called neuroeconomics.⁴² This literature maintained the normative – descriptive distinction, but nevertheless slightly re-interpreted the distinction by supposing that the two sides of the distinction represent two sides of human behavior. In other words, the normative was reduced from something external to the individual to one of two faculties innate to human nature that strive for dominance. An illustrative example is McClure, Laibson, Loewenstein, and Cohen (2004), "Separate Neural Systems Value Immediate and Delayed Monetary Rewards." The research described in the article sought and found evidence for neurobiological substrates for the two components of the effective discount factor as described above. When faced with delayed monetary rewards while lying in an MRI-scanner, subjects' brains showed peaks of activity in the parts of the brain associated with rational behavior (in this case the lateral prefrontal cortex and the posterior parietal cortex); when faced with immediate rewards, the limbic system associated with the midbrain dopamine center was especially active.

⁴² The label neuroeconomics is used by at least two groups of scientists [see e.g. Vromen (2007), Ross (2008)].

The authors took their findings as evidence that, when faced with the choice between an immediate reward and a higher future (thus ‘expected’) reward, the two parts of the brain strive for dominance. The limbic system was especially sensitive to immediate rewards and signaled choosing the immediate reward. The prefrontal cortex was more sensitive to the higher expected pay-off and signaled choosing the delayed reward. In other words, the experimental results were taken as evidence for the postulated difference between the short-term hyperbolic discounting, and the long-term rational exponential discounting. The normative long run rational system strives for dominance with the short term affective system. When the short run system is affecting the outcome, the resulting behavior will be observed and classified as systematically deviating from the norm. As such, the neuroeconomic research conducted by behavioral economists at once maintained the link with the normative – descriptive core of Kahneman and Tversky’s work, while at the same time constructing a link with neuroscience.

Based on this and other research, in the 2000s behavioral economists have increasingly argued that the neuroscientific framework should be adopted as a basis for investigating individual (economic) behavior. The recurring argument in the neuroeconomic research has been that in economic decision making the individual’s rational system tries to make the rational decision, but, alas, is often overridden by a strong and dominant affective system. Intertemporal choice has provided a good example. When people need to plan how to divide their income between consumption and spending at some point in the future, their affective system is not very much involved and the rational system will decide on a rational division between the two. However, when the future becomes the present, the affective system kicks in, seeking immediate gratification and thereby seeking to override the rational system. In other words, the systematic deviations from rational decision making were understood as having resulted from a failed attempt by the individual’s rational system to control its affective system.

Within this general neuroeconomic approach behavioral economists have proposed different frameworks for the two systems. Let me give two prominent examples. In the paper that derived from his Nobel lecture [Kahneman (2003)], Kahneman argued for and employed the following framework.

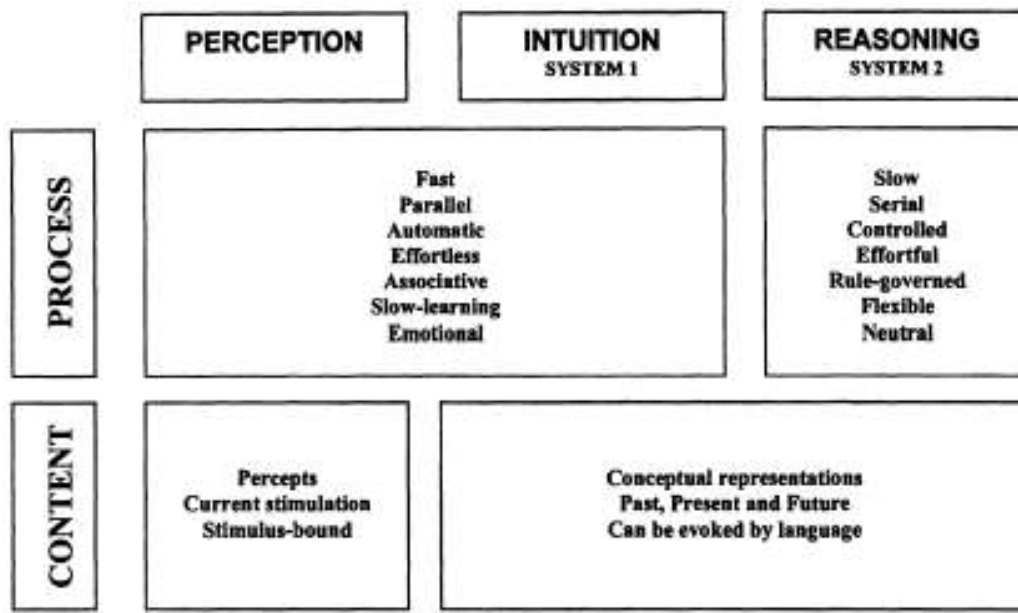


FIGURE 1. THREE COGNITIVE SYSTEMS

Source: Kahneman (2003), p.1451

Kahneman's framework is an intriguing mix of psychophysics, neuroscience, and the desire to accommodate a distinction between two decision making processes. On the bottom row we see a distinction between two kinds of input for decision making: current external stimulation and information already present in the mind. These two sources of information form the input for two cognitive systems as described in the middle row: an intuitive system (also more neutrally labeled system 1), and a reasoning system (also more neutrally called system 2). The distinction between the two systems is mainly made in terms of the effort it costs to operate them. The reasoning system requires much effort and is relatively slow. The intuitive system operates much more quickly and is relatively effortless. Another distinction between the two is that between non-voluntary impressions in system 1, and voluntary judgments in system 2. The top row further distinguishes between perception and the intuitive information processing of system 1. The distinction between the two information processing systems and perception is made in terms of automaticity and accessibility. On the left we find decision making that is fully automatic and inaccessible. One example is the perception/decision regarding which of two rooms has a higher temperature, a perception/decision that is made automatically and without the individual having access to its process. At the other end of the spectrum, we find decisions that are made non-automatically and to which a large degree of

accessibility is possible. An example is the decision regarding which of two houses is preferred and hence will be bought. The intuitive system is a middle ground between these two and reflects decision making that often proceeds automatically, but which can also be accessed and altered, such as the decision between € 3,000 for certain and a 0.8 chance at € 4,000. Using this framework Kahneman thus brought together the automatic perceptual system investigated by psychophysics and behavioral economics' use of neuroscience, while at the same time allowing for the possibility of deviations from rational economic decision making.

Another prominent framework, extensively discussed in what may safely be considered as one of the canonical articles in neuroeconomics, Camerer, Loewenstein, and Prelec (2005), "Neuroeconomics: How Neuroscience Can Inform Economics" is the following:

TABLE 1 TWO DIMENSIONS OF NEURAL FUNCTIONING		
	Cognitive	Affective
Controlled Processes		
<ul style="list-style-type: none"> ■ serial ■ effortful ■ evoked deliberately ■ good introspective access 	I	II
Automatic Processes		
<ul style="list-style-type: none"> ■ parallel ■ effortless ■ reflexive ■ no introspective access 	III	IV

Source: Camerer, Loewenstein and Prelec (2005), p.16

In this representation the two systems are either controlled or automatic. Examples of automatic, cognitive processes, quadrant III, are judgments of relative temperature and shapes of objects. The automatic affective system, quadrant IV, depicts the pleasure and pain system that on the basis of the information provided by quadrant III and on the basis of information of past experiences attaches a value to the object. Quadrant I and II constitute the cognitive and affective part of the decision making process that can be controlled. A decision-maker may very much like to buy a car, but in quadrant I reason determines that he or she cannot afford it. And he or she may not

be hungry at all, while in quadrant II, not wanting to disappoint his or her friend's cooking efforts, he or she takes a bite nevertheless.

Thus, in behavioral economics and its neuroeconomics research, descriptive real-world decision making should be seen as a struggle between a cognitive system that seeks a rational solution and an affective system that disregards the optimal decision in favor of immediate gratification. In that respect, the dual system approach of contemporary behavioral economics is different from the thinking of many other neuroscientists and social scientists who do not see the cognitive system as superior to the intuitive, emotional, or affective system. Examples of the latter include the work of Damasio [e.g. Damasio (2003)], research descending from Simon [e.g. Gigerenzer et al. (1999), Gigerenzer and Selten (2001)], and evolutionary theory inspired science [e.g. Barkow, Cosmides and Tooby (1992)].

The intertemporal choice and dual systems approach literature illustrates that in building behavioral economics, behavioral economists have incorporated the normative-descriptive distinction in their theories of economic decision behavior. They have come to see the normative not as something external to the individual, but as a rational system side by side an affective system, with which it strives for dominance. As Kahneman has also contributed to this new and changed perception of the normative and descriptive, it should be seen as reflecting a development in the behavioral decision research community at large, not only a development in behavioral economics.

2.2 Are preferences innate to human nature or do they emerge through interaction with the environment?

The incorporation of the distinction made between the rational norm and its imperfect realization in the economic agent considerably broadened the scope of behavioral economics. However, there seemed to be limits to broadening the scope as well. This can be illustrated by behavioral economists' cooperation with anthropologists on the subject of the emergence of preference. This brief collaboration shows how behavioral economists, after their initial enthusiasm retreated when it turned out that this collaboration resulted in research that was at odds with the fundamental behavioral economic assumption of a fixed, universal benchmark of full rationality. As in the case of intertemporal choice, behavioral economists were unwilling to give up this benchmark of full rationality. The research was a large scale interdisciplinary

study of the ultimatum-game in fifteen small-scale societies. It was published in a number of journals. The most extensive discussion can be found in the book devoted to it, Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr and Herbert Gintis, *Foundations of Human Sociality* (2004). A reflection on the research summarized by Henrich can be found in Gigerenzer and Selten (2001).⁴³

The motivation for this large interdisciplinary study was the following. The ultimatum game had been played all over the world, and has always led to the result that individuals do not play the rational optimum, but typically divide the money about half-half.⁴⁴ However, the experiments had only been done with university students in advanced capitalist economies. The question was thus whether the results would hold up when tested in other environments.

The surprising result was not so much that the average proposed and accepted divisions in the small-scale societies differed from those of university students, but how they differed. Roughly, the average proposed and accepted divisions went from [80%,20%] to [40%,60%]. Individuals in different societies thus showed a remarkable difference in the division they proposed and accepted. Henrich and his fellow researchers correlated these differences with two economic characteristics of small-scale societies. First, they documented how much a group's (normally the family) economic welfare depended on co-operation with other groups within a small-scale society. In this respect the societies differed greatly from almost none to almost completely. Second, the researchers investigated how much the group's economic welfare depended on market exchange. There were also differences in the level of market integration. The researchers concluded that differences found in the behavior of individuals belonging to the different societies in the game should be attributed to differences in the environment in which they lived. As a consequence, preferences were not exogenous, but determined by the environment. Henrich and his collaborators stated this explicitly in a brief summary of their research in the *The American Economic Review*:

⁴³ The research was funded through the MacArthur Foundation's research network *The Nature and Origin of Preference*. The network, "headed by Robert Boyd and Herbert Gintis received two grants: \$2.55 million in 1997 and \$1.8 million in 2002" [email Richards to author, December 4, 2008].

⁴⁴ The ultimatum game is the following: Player one proposes a division of a fixed sum of money, player two either accepts (the money is divided according to the proposed division), or rejects (both players get nothing). If both players are rational player, one proposes the smallest amount possible and player two accepts.

[...] preferences over economic choices are not exogenous as the canonical model would have it, but rather are shaped by the economic and social interactions of everyday life. This result implies that judgments in welfare economics that assume exogenous preferences are questionable, as are predictions of the effects of changing economic policies and institutions that fail to take account of behavioral change. [Henrich et.al. (2001), p.77]

Giving up the exogeneity of preferences would have far reaching implications. It would mean that behavioral economists had to give up a fundamental tenet of economics, namely that preferences are given. That would have in turn far-reaching consequences for the theory of decision making itself. It would draw behavioral economists closer to cultural anthropologists, such as Boyd or Henrich, who conceive culture in evolutionary or ecological terms. Taking this approach and the results of the experiments seriously would imply not only that individual preferences to a large extent are determined by the environment and by learning, but it would also undermine the notion of fixed norms in behavioral decision research and well-defined rationality in economics. It is probably due to these extensive implications that in spite of plans to continue the research, follow-up studies and further elaboration of the implications of the experiments by this diverse group of economists and anthropologists to date have not been worked out.

2.3 Understanding behavioral economics in terms of rationality

When behavioral economics expanded, behavioral economists were both faithful to the Kahneman and Tversky legacy, while at the same time they sought to broaden its scope. Problematic in this regard were the labels of normative and descriptive, which were considered confusing in an economic context that already had created its own understanding of these concepts. We have seen examples of how economists changed this distinction by developing different models of economics decision making. As a consequence, behavioral economists in the 1980s-2000s have reinterpreted the normative-descriptive distinction in terms of rationality.

As set out in Chapter five, Thaler was well aware of the fact that the re-interpretation of economics in terms of normative versus descriptive raised the question concerning the definition of the descriptive theory when the normative theory is about rational behavior. And here Thaler was not very specific, or at least he

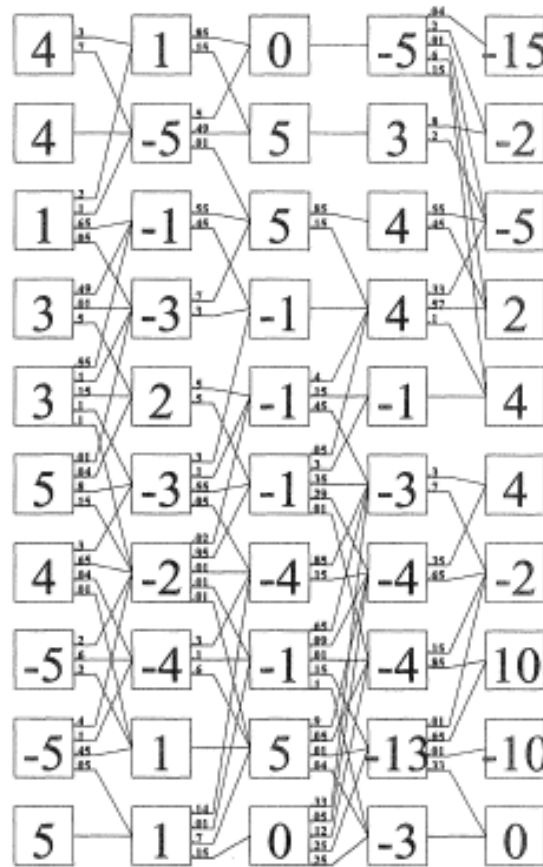
did not offer a conclusive answer. Thaler referred to behavior that deviates from the normative solution on a number of occasions as “irrational” or “non-rational.” Furthermore, he noted that he “would not want to call such choices rational” [Thaler (2000), p.138]. On other occasions Thaler referred to the normative-descriptive distinction as rational versus emotional [see e.g. Shefrin and Thaler (1988), p.611].

But the main interpretation Thaler used in the 1980s and 1990s was the term “quasi rationality,” most prominently as the title of a collection of articles, *Quasi Rational Economics* (1991). Quasi rationality suggests a category of behavior somewhere between the full rationality of the normative decision and irrational behavior. Regularly used in the 1980s and 1990s quasi rationality is perhaps best understood as the failed attempt of people to be rational, which is exemplified by the one suggested definition of the term that Thaler has provided: “quasi rational, meaning trying hard but subject to systematic error” [Thaler (2000), p.136]. On another occasion it was characterized as “less than fully rational” [Thaler (1991), p.xviii].

From the early 2000s onwards, the term which has been increasingly favored by behavioral economists is “bounded rationality.” The distinction that was made is that between the fully rational decision and the decision actually made that has been deemed boundedly rational when deviating from the rational decision. Full rationality in behavioral economics was defined as follows:

The standard approach in economics assumes “full rationality.” While disagreement exists as to what exactly full rationality encompasses, most economists would agree on the following basic components. First people have well-defined preferences (or goals) and make decisions to maximize those preferences. Second, those preferences accurately reflect (to the best of the person’s knowledge) the true costs and benefits of the available options. Third, in situations that involve uncertainty, people have well-formed beliefs about how uncertainty will resolve itself, and when new information becomes available, they update their beliefs using Bayes’s law – the presumed ability to update probabilistic assessments in light of new information. [Camerer et al. (2003), 12-14-1215)]

Using the distinction between full and bounded rationality naturally entailed making references to the work of Herbert Simon. This can be illustrated by Gabaix and Laibson's article "A Boundedly Rational Decision Algorithm" (2000). In this article, Gabaix and Laibson created decision algorithms for a specific decision problem that deviated from the "fully rational" algorithm, to which the term "boundedly rational" was applied. The following simulation was constructed. Consider a five-level decision tree where at each level, including the first, there are ten different pay-off boxes and between one and five connections to the next level. Each of the pay-off boxes' possible connections to the next level has a certain probability, with the added probability of the connections of one box always amounting to one. An example of such a randomly generated decision tree appears as follows.



Source: Gabaix and Laibson (2000), p.434

Gabaix and Laibson argued that for a fully rational decision algorithm, the decision maker would be required to calculate the expected pay-off of each box and then make

a decision about which of the ten first level boxes to start. However, this would require a considerable amount of computational capacities and time. The authors accepted from Simon's work that both are often unavailable. They considered three "boundedly rational" alternative algorithms: i) "Follow-the-leaders" (FTL), which ignores probability paths of less than 0.25, ii) column cut-off, which ignores one or more of the last columns, and iii) discounting, which discounts the values of later columns. Of these three alternatives, FTL came closest to the fully rational outcome. Furthermore, it closely matched behavior of subjects in an experiment who were faced with the same task. It was concluded that individuals' behavior might be explained in terms of a boundedly rational algorithm.

In other words, the concept of bounded rationality was taken from Simon and together with the concept of full rationality employed to rephrase Kahneman and Tversky's normative – descriptive distinction. In one clear sweep, Kahneman and Tversky's distinction between the concepts of normative-descriptive were replaced by concepts more appropriate in an economic context, and at the same time Simon was appropriated as an authoritative source for the use of these concepts. Simon provided behavioral economics with the fitting language and offered it the necessary authority. This is exemplified by Kahneman in the *American Economic Review* article based on his Nobel lecture, entitled "Maps of Bounded Rationality: Psychology for Behavioral Economics" (2003). In the opening passage Kahneman used the term "bounded rationality" and referred to Simon, but at the same time subtly but clearly distinguished his and Tversky's work from Simon's.

The work cited by the Nobel committee was done jointly with Amos Tversky (1937-1996) during a long and unusually close collaboration. Together, we explored the psychology of intuitive beliefs and choices and examined their bounded rationality. Herbert A. Simon (1955, 1979) had proposed much earlier that decision makers should be viewed as boundedly rational, and had offered a model in which utility maximization was replaced by satisficing. Our research [on the contrary – FH] attempted to obtain a map of bounded rationality, by exploring the systematic biases that separate the beliefs that people have and the choices they make from the optimal beliefs and choices assumed in rational-agent models. [Kahneman (2003), p.1449]

In a clever way, Kahneman invoked Simon to construct authority for the behavioral economic program, while at the same time interpreting the concept of bounded rationality in such a way that it would become fully compatible with his and Tversky's approach and that of the behavioral economists. Thus, at the same time behavioral economists remained committed to Kahneman and Tversky's methodological distinction between separate normative and descriptive domain, re-labeling this distinction in more appropriate economics terms, and claiming Simon's intellectual heritage as adding credibility to their research.

2.4 Economic policy as a form of paternalism

The re-interpretation of Kahneman and Tversky's distinction between the normative and the descriptive in terms of a conflict within the economic decision maker had important consequences for welfare economics. Most mainstream economists in the 1990s and 2000s associated welfare economics in one way or another with the term normative. That was one reason why Kahneman and Tversky's labels of normative and descriptive invoked confusion when inserted into economics discourse. The reinterpretation of normative versus descriptive in terms of full-rationality versus bounded rationality solved this confusion and in turn allowed behavioral economists to develop their own position on welfare economics. Behavioral economists described their new stance on welfare issues as a form of paternalism. Behavioral economic paternalism, then, resulted from the solution that behavioral economists found in dealing with the violations of the normative or full-rationality benchmark. It illustrates the ultimate consequence behavioral economists have drawn from their new program.

The paternalism debate involves the application of behavioral economic insights into real-world policy questions. For instance, behavioral economics had discovered that people often save much less for their pensions than they should, and that when they do save, they do not diversify their portfolios optimally. Following on these results, programs have been set up to investigate how people can be induced to save more for retirement and better diversify their stock portfolios [e.g. Cronqvist and Thaler (2004), Thaler and Benartzi (2004)].

Another example concerns the use of medication. It has often been found that people who need to take drugs on a regular basis are very lax at doing so. Even when the risks are substantial and potential costs in terms of health very great, such as in the case of medication that reduces the chance of having a second stroke, people are very

lax at taking their medication properly. To solve this problem, programs have been set up that investigate how insights from behavioral economics can be used to design incentive mechanisms that induce people to take their medication [e.g. Badger et al. (2007)]. Finally, behavioral economists have turned their attention to development economics, with the purpose of using insights from behavioral economics to improve the functioning of development programs [see e.g. Mullainathan (forthcoming), and Betrand, Mullainathan and Shafir (forthcoming)]

Behavioral economists have framed and defended this research in a number of closely related ways. Well-known is Thaler and Sunstein's (2003) "Libertarian Paternalism." Libertarian paternalism can be understood as a paternalism that does not restrict individual freedom of choice. Thaler and Sunstein distinguish themselves explicitly from the Samuelsonian stance towards welfare issues.

We clearly do not always equate revealed preference with welfare. That is, we emphasize the possibility that in some cases individuals make inferior choices, choices that they would change if they had complete information, unlimited cognitive abilities, and no lack of willpower.

[Thaler and Sunstein (2003), p.175]

In the behavioral economics paternalism debate, the justification for paternalistic policies has been the fact that decisions people actually make, their "revealed preferences," do not always match with their "true" preferences. Behavioral economists have thus constructed a distinction between "revealed" and "true" preferences. This does not mean that preferences are context dependent. Rather, it means that it depends on the context whether the true preferences can and will be revealed appropriately. A source that is sometimes relied on in this regard is John C. Harsanyi who noted that "in deciding what is good and what is bad for an individual, the ultimate criterion can only be his own wants and his own preferences," where the individual's "own preferences" are his "true" preferences: "the preferences he would have if he had all the relevant factual information, always reasoned with the greatest possible care, and was in a state of mind most conducive to rational choice" [quoted in Angner and Loewenstein (forthcoming), pp.53-54].

A more detailed and elaborate explication and defense of this new branch of behavioral economics can be found in Camerer et al. (2003) "Regulation for

Conservatives: Behavioral Economics and the Case for ‘Asymmetric Paternalism’.” In this article, the five authors (Camerer, Issacharoff, Loewenstein, O’Donoghue and Rabin) make a case for what they label “asymmetric paternalism,” where “[a] regulation is asymmetrically paternalistic if it creates large benefits for those who make errors, while imposing little or no harm on those who are fully rational” [Camerer et al. (2003), p.1212]. Behavioral economics, then, “describes ways people sometimes fail to behave in their own best interests” [Camerer et al. (2003), p.1217]. These “apparent violations of rationality [...] can justify the need for paternalistic policies to help people make better decisions and come closer to behaving in their own best interests” [Camerer et al. (2003), p.1218].

Thaler and Sunstein (2003) countered possible aversions to paternalism by economists and others by linking paternalism to libertarianism. Camerer et al. (2003), on the other hand, founded their defense of paternalistic policies on the need for asymmetry in paternalistic policy. The definition of asymmetric paternalism resembles the Paretean improvement argument: “a policy is *asymmetrically paternalistic* if it creates large benefits for those people who are boundedly rational [...] while imposing little or no harm on those who are fully rational [Camerer et al. (2003), p.1219, emphasis in the original]. Or, in other words, “asymmetric paternalism helps those whose rationality is bounded from making a costly mistake and harms more rational folks very little” [Camerer et al. (2003), p.1254]. Another way of putting it, Camerer et al. (2003) argued, is to see the limitedly rational individual as imposing negative externalities on his or her own demand curve. “When consumers make errors, it is as if they are imposing externalities on themselves because the decisions they make as reflected by their demand do not accurately reflect the benefits they derive” [Camerer et al. (2003), p.1221]. Hence, there is a need for a policy maker who can remove the externalities and redirect behavior in such a way that the externalities disappear. Camerer et al. (2003) furthermore noted that firms could either consciously or unconsciously use the irrationality of individuals to gain more profit.

A detailed example has been provided by Grubb (2007), who showed that cell-phone companies could permanently increase their profits using the phenomenon of overconfidence. When consumers systematically underestimate the number of minutes they will use their cell phones each month, it is profitable for firms to charge the marginal costs per minute up to the number consumers expect they will use their

cell-phones. After this point, they can greatly increase their rates. Consumers will not mind because they do not expect to be using their phones that much. But the cell-phone company knows that the consumers are overconfident, and thus knows the customers will use some of these expensive minutes. Thus the cell-phone company increases its profit by exploiting consumers' overconfidence. The example shows that firms in the market can use the limited rationality of individuals and thus *not* produce efficient outcomes. Moreover, the example shows that it might be more in the interest of companies to maintain or even amplify the limited rationality of individuals. It suggests that firms are apt to look for ways to increase the irrationality of individuals, and that as a result the market, instead of producing an efficient equilibrium, could even produce greater irrationality. Behavioral economists have shown that deviations from full rationality do not only persist in markets, but that markets can even increase this deviating behavior.

On the basis of these results, behavioral economists have argued that economists are morally obliged to act against the violations of full rationality:

As economists, how should we respond to the seemingly self-destructive side of human behavior? We can deny it, and assume as an axiom of faith that people can be relied upon to do what's best for themselves. We can assume that families paying an average of \$1,000 per year financing credit card debt are making a rational tradeoff of present and future utility, that liquidity constraints prevent investing in employer-matched 401k plans, that employees prefer investing in their own company's stock instead of a diversified portfolio [...] that people are obese because they have calculated that the pleasures from the extra food, or the pain of the foregone exercise, is sufficient to compensate for the negative consequences of obesity. [Loewenstein and Haisley (forthcoming), p.4]

According to behavioral economists, economics is particularly suited for solving the violations of full rationality because it possesses the knowledge of how to "steer human behavior in more beneficial directions while minimizing coercion, maximizing individual autonomy, and maximizing autonomy to the greatest extent possible" [Loewenstein and Haisley (forthcoming), p.6]. The role of the economist in this regard can be seen as analogous to the psychoanalytical therapist. "Just as the

therapist endeavors to correct for cognitive and emotional disturbances that detract from the well-being of the patient, such as anxiety, depression, or psychosis, the economist/therapist endeavors to counteract cognitive and emotional barriers to the pursuit of genuine self-interest” [Loewenstein and Haisley (forthcoming), pp.9-10].

Behavioral economists have attempted to solve mankind’s limited rationality problem by using phenomena similar to those that formed the basis for behavioral economics to begin with. The reason that they could do so was that behavioral economics had remained faithful to Kahneman and Tversky’s approach. The most important phenomenon in this regard is what is most commonly known in behavioral economics as framing. One of the central findings of Kahneman and Tversky’s behavioral decision research and behavioral economics was that people are susceptible to the way in which a choice is presented to them. Depending on the “reference point,” in Kahneman and Tversky’s terms, or “frame,” the term Thaler favored for behavioral economics, people change their preferences. The example taken from Thaler and Sunstein (2003) is of the cafeteria manager who can either place the desserts before the fruits or vice versa. If he or she frames this decision as fruits-before-desserts, the fruit will be chosen more often. Thus, framing is used to influence people’s behavior without affecting their freedom to choose in any significant way. Changing the default option from not-participating to participating in pension saving schemes is another often-quoted example.

By exploring how policies can be designed to solve the bounded rationality of individuals, behavioral economists have taken the full-rationality versus bounded rationality framework and the experimental results of psychologists and economists to their ultimate consequences. Behavioral economic paternalism is very an economics solution to bounded rationality, emphasizing incentive mechanisms and monetary rewards.

3. Defining behavioral economics

The different explorations of behavioral economists in the 1990s and 2000s contributed to the gradual construction of a clearly defined field. Over the course of some twenty years, behavioral economics used the conceptual core as laid out by Kahneman and Tversky and explored its implications in rethinking the economic agent. Rather than the received rational decision maker of neoclassical theory, the individual became a decision maker in whom the cognitive and the affective system

were in competition. Despite the fact that behavioral economists questioned the normative benchmark at different points in time, in the end this distinction remained a core element of behavioral economics, though in a modified form. Over time the labels were re-interpreted in terms of full and bounded rationality. In turn, this allowed behavioral economists to claim the heritage of Simon, though they took a different perspective than other social scientists who claimed to work in the tradition of Simon, namely what could be broadly and roughly labeled the ecological approach to rationality. This group includes anthropologists such as Boyd and Henrich, psychologists such as Gigerenzer, and experimental economists such as Vernon Smith and Charles Plott. But behavioral economists' use of Simon also made them more clearly aware of how they differed from other strands involved with decision research.

These theoretical and conceptual developments formed only one aspect of a gradual process which defined behavioral economics. In addition to this theoretical and conceptual definition, behavioral economists came to define themselves across disciplinary lines. In particular, they gradually began to distinguish behavioral economics from the discipline of psychology and from the sub-discipline of experimental economics. In the case of psychology, the reason was that behavioral economists wanted to be accepted by the economic mainstream. To achieve this they had to comply with the mainstream economic view that economics and psychology are separate disciplines and that economics is superior to psychology when it comes to rigor and formal modeling. Some aspects regarding the attempts made by behavioral economists to distinguish themselves from psychology have been hinted at above in the discussion concerning Laibson's way of incorporating psychological insights. More specifically, behavioral economists have employed the following arguments to distinguish behavioral economics from the discipline of psychology.

First, behavioral economics has been defined as economics on the basis of its use of mathematical modeling. This argument has remained largely implicit, as in the case of Laibson, but it is nevertheless clear that behavioral economists consider their use of mathematics superior to that of the psychologists, and that they see it as a defining characteristic of economics. Another illustrative example in this regard is the work of Matthew Rabin [e.g. Rabin (1993, 1994, 1998)]. Rabin has incorporated experimental results from psychology into economics, but has combined this with advanced economic mathematical modeling. He argues that "none of the broad-stroke arguments for inattention to psychological research are compelling," but at the same

time cautions that “not all psychological research will be [...] proven to be of great economic importance.” The reason is that because they come from psychology, these psychological results have not yet been subjected to “the same rigorous standards that our discipline, at its best, applies elsewhere” [Rabin (1998), p.41]. That is, ultimately it requires the mathematics of economics to judge how useful the insights from psychology are. This use of mathematics is something that defines behavioral economics as economics, and therefore as different from psychology.

Second, behavioral economists distinguished themselves from psychology on the basis of their use of the experimental method. For instance, in a methodological comment on sharing data or experiment instructions with other researchers, Camerer noted that “[i]f you asked a psychologist for data or instructions he or she might be insulted, because the convention in that field is to give the writer the benefit of the doubt” [Camerer (2003), pp.34-35].⁴⁵ Another example concerns the use of deception. A common experimental procedure in psychology is to tell the experimental subjects the experiment is about one thing, while it is in fact about something else. Behavioral economists and experimental economists have resisted this method of deception ever since experiments have been used in economics. The reason for instance that Grether and Plott (1979) controlled for the fact that the experimenters had been psychologists was that they suspected that because the subjects knew they might be misled in a psychological experiment they would behave differently. In recent years the use of deception has produced increasing discussion in the experimental communities of economics and psychology [e.g. Hertwig and Ortmann (2008), Ortmann and Hertwig (2002), Jamisona, Karlaub, and Schechter (2008)].

The motivation for explicitly drawing a border between behavioral and experimental economics was different. Both research programs were part of the same economic science, and to define behavioral economics as economics, it was not necessary to distinguish between behavioral and experimental economics. The reason that behavioral economists nevertheless wanted to distinguish their field explicitly was that they had come to realize that theoretically they were of a fundamentally different opinion than experimental economists regarding the relation of individual decision making to market outcomes. However, because experimental economists used the same experimental method as behavioral economists, outsiders could easily

⁴⁵ Interesting in this regard is that Camerer holds a PhD (defended in 1981) in behavioral decision research, and not in economics. His supervisor at the University of Chicago was Robyn Hogarth.

conclude that they must be the same. Therefore, the two needed to be clearly distinguished. The distinction was drawn on the basis of the experimental method.

The different views of experimental economics by behavioral economists have been grouped together in Loewenstein (1999) “Experimental Economics from the vantage-point of Behavioural Economics.” Loewenstein (1999) positioned behavioral economics explicitly in opposition to experimental economics. He formulated his critique in terms of the “psychological distinction” of external versus internal validity. Under the heading of external validity, Loewenstein saw four problems with experimental economics. First, experimental economics put great emphasis on the use of auctions in its experiments. As people in reality hardly ever find themselves in an auction situation, it is doubtful that these experiments can tell us very much about economic behavior in the real world. Second, Loewenstein disagreed with experimental economists’ use of repetition in what has been called the Ground Hog Day argument. In reality, Loewenstein argued, people never make the exact same decision forty times in a row. Real world behavior is much more like the first few rounds of an experiment than the last two or three rounds. Third, Loewenstein criticized experimental economists for their tendency to reduce real-world content to the absolute minimum possible. Apart from the fact that a context-free experiment is an illusion, Loewenstein argued it also greatly reduces the external validity of the experiments.⁴⁶ Instead, economists should, just as Loewenstein himself, make the experimental situation as congruent with reality as possible; hence make the experiment “context-rich.” Fourth, according to Loewenstein experimental economists wrongly assumed that monetary rewards result in strict control over incentives. It has been shown in numerous experiments that this is not the case. With monetary incentives, subjects are also likely to be driven by other motives than profit maximization. Finally, one problem concerning internal validity that Loewenstein observed was that experimental economists had been far too careless in not using randomization and in comparing the experimental results that had been obtained under different circumstances.

⁴⁶ Loewenstein both uses ‘context’ and ‘content.’

4. Behavioral economics in the 2000s: stable, defined, but slightly schizophrenic

In the 1990s and 2000s behavioral economics was built, defined, and distinguished as a stable mainstream economic program. It was built on and remained true to the framework as laid out by Kahneman and Tversky, although it re-labeled its key concepts in terms of rationality. On the basis of their explorations, behavioral economists came to define step by step behavioral economics both theoretically and conceptually as an economic program that seeks to develop theories of bounded rationality in order to complement the normative benchmark of full rationality. In turn, this has laid the foundation for a new approach to welfare policy under the rubric of paternalism.

But the emerging behavioral economics has also been somewhat schizophrenic. Behavioral economists have defined behavioral economics as an important improvement on the traditional neoclassical economic paradigm by this incorporation of insights from psychology, but at the same time they have sought to define behavioral economics as being different from psychology. They have claimed the heritage of Simon and his bounded rationality, but at the same time they have explicitly distinguished behavioral economics from the work of other social scientists using Simon's work. Yet, they have not done so by discussing how Simon should be interpreted, but instead on the basis of their view of the proper way to conduct experiments in economics. And finally, they have sought to define behavioral economics as mainstream economics, but at the same time they have gone to great lengths to embed behavioral economics more firmly in the cognitive and neuro sciences.

The reason for this slight schizophrenia is that the incorporation of psychology has not been a neutral process. In fact, behavioral economists have used psychology to redefine economics. The psychological theories, experimental results and authoritative figures have not been used for their own sake but solely because they could be used to steer economics in a different direction. Part of this new direction has been a language of "enriching" economics with psychology, but this is merely a rhetorical guise for convincing economists which part of psychology they should understand, and in which way that should alter economic reasoning. That behavioral economists have not been neutral receivers of psychological input, but have actively built, defined and distinguished behavioral economics using psychology does not

diminish or undermine the behavioral economic project. However, it does raise the question regarding how we ought to understand the flow of theories, methods and results that cross from one discipline to the other.

7. Changing meanings: disunity in economics and psychology

In the first chapter, I introduced Peter Galison's notion of disunity and applied it to economics and psychology. Economics and psychology in the postwar period were disunified cultures. The two disciplines communicated with one another, and theories, methods and experimental results travelled between them. But the communication and exchange were not understood in the same way. By and large, psychologists understood economics to be only interested in the normative aspects of economic behavior. As such it was understood to be a sub-discipline of psychology. For economists, on the other hand, psychology and psychological assumptions constituted the necessary premise upon which economic theories and models were constructed. Economics started where psychology left off, just as chemistry starts where physics ends.

This thesis has shown how from the late 1970s and early 1980s onwards economists have used psychology to redefine economics. In this period, behavioral decision researchers communicated with experimental economists and with financial economists, and as a result theories, methods, and experimental results travelled from psychology to economics. This was not a neutral exchange that left theories, methods and experimental results unaffected. Instead, the meaning of the theories, methods, and experimental results exchanged altered when they entered economics. They lost some of their psychological connotations and gained new economic connotations. For instance, the experimental method and experimental results lost their role of testing the human measurement instrument, and continued as only testing human behavior against the normative benchmark of rationality. Prospect theory lost its connotation of unifying psychology and economics in one field of behavioral science, although this took several years before it became clear.

What is particularly illustrative in this regard are the two very different ways in which psychological theories, methods, and experimental results were interpreted in experimental economics and financial economics. The theories, methods and experimental results that travelled from psychology to economics acquired very different new meanings in these two economic programs. In addition to the disunity between psychology and economics, a disunity appeared in the early 1980s between experimental and behavioral economics.

In experimental economics, the corroboration of psychologists' experimental results was interpreted as completely invalidating any description of individual behavior as rationally maximizing utility. Therefore, in experimental economics prospect theory was interpreted as showing how even an adjusted rational maximization theory was not descriptive of individual behavior. In addition, the experimental results were interpreted as proving the importance of the market as a rationalizing mechanism that over time drives the economy to equilibrium. The experimental results lost their relation to the traditional normative benchmark, but acquired a new meaning by demonstrating a phenomenon unexamined by psychologists: the market. In contrast, in behavioral economics the experimental results of the psychologists maintained their meaning in disproving the normative theory of rational decision making. However, this was merely the first step in producing new additional meanings they were to be given in behavioral economics, namely as invalidating the norms of rationality in markets. In behavioral economics, psychologists' experimental results were interpreted as disproving the existence of rational equilibria in markets, and prospect theory was interpreted as a first and promising candidate for providing a descriptive alternative that could explain both deviating individual behavior and irrational markets.

As a consequence, these different cultures and sub-cultures also employed and understood the experimental method in different ways. For psychologists, the experimental method was a way to investigate the human being in its natural state. It was a method which could be used to investigate the underlying characteristics of human behavior that determine or affect all human behavior, irrespective of its particular background, knowledge, or cultural specifics. For example, in order to reduce the noise from particularities of the individuals in the experiments as much as possible deception was often considered a valuable method. It ensured that experimental subjects would not respond on the basis of prior knowledge of the theory behind an experiment, specific cultural beliefs of proper behavior and so forth; a methodological problem in psychology generally known as the Hawthorne effect [e.g. Adair (1984)]. In contrast, both experimental economists and behavioral economists ruled out deception because they used the experimental method to investigate how the individual behaved in different economic situations. It was not the individual's underlying nature, but his or her economic behavior that was the subject of investigation. Therefore, it was vital that the subjects in the experiments completely

understood what situation they were in, and that they never felt as if they were being fooled by the experimenter. The individual should behave as if he or she was in the economic environment simulated by the experimental set up, and should not behave as an individual who does not completely understand what is going on and/or suspects he or she is being fooled in some way or another. Thus, the experimental method had a different meaning in psychology and economics.

One difference between the use of experiments in experimental and behavioral economics was that for experimental economists experiments were a method of testing the functioning of the market that operates over and above individuals, whereas for the behavioral economists experiments tested market behavior as a characteristic of individual behavior in a market environment. For behavioral economists, market or economic behavior was about the individual, and the testing of this behavior only required that individuals be given an economic choice. For experimental economists, on the other hand, the market was a mechanism, an invisible hand, that over and above individual behavior directed the behavior of each economic agent towards achieving more rationality. This market could only be tested through individuals, who in the controlled setting of an experiment behaved as if they were in a real economic market. Because of this different understanding of the market, experimental and behavioral economists also employed experiments differently. Behavioral economists employed experiments to carefully observe how individuals respond to specific economic choices. Experimental economists, on the other hand, employed experiments to investigate how individual economic behavior adapts over time when individuals face the same situation a large number of times, but in each round acquire more information about the behavior of the other participants and more information about the functioning of the market. Thus, in addition to the different meanings of the experimental method in psychology and economics, the experimental method acquired a different meaning in the two subcultures of experimental and behavioral economics.

The disunity of psychology and economics is further illustrated by their different understanding and use of measurement. To behavioral decision researchers, measurement and individual behavior were two sides of the same coin. Coming from the German psychophysical tradition, the theory of human decision making was at once understood as a theory of human perception and rational inference, and as a theory of the human being as a measurement instrument. In psychological

experiments, the underlying nature of human decision behavior as well as the human measurement instrument was investigated. But when psychological theories, experimental methods and experimental results travelled to economics, this double meaning was lost. In both behavioral and experimental economics discussions of measurement were simply absent.

All these different elements of the disunity of economics and psychology and the disunity of experimental and behavioral economics explain the different ways economists and psychologists have understood the relationship between their disciplines. The psychologists saw the economists using the same method and same theory as used in psychology and thus considered economists to be engaged in the same project as psychology. Because of this understanding of economics, Kahneman and Tversky advanced prospect theory as a theory meant to unify the two disciplines that for some unknown reason had been developing separately alongside each other. Economists on the other hand saw psychology as investigating human behavior, and therefore as a possible source of information for economics' starting point. The difference between experimental and behavioral economics was that for behavioral economists changing the behavioral assumptions had direct consequences for the economic models and theories, whereas for experimental economists it did not. Economics and psychology were disunified in their understanding of the relationship between economics and psychology, but experimental and behavioral economists also had different understandings of the relationship between the two. The difficulty in understanding the history of the mutual influence of economics and psychology is not only that influencing factors meant different things in the different disciplines, but that as a consequence also what the different disciplines were about and how they were related were understood differently. Across the different (sub-)disciplines and across time, the definitions of economics and psychology have not been stable entities.

Finally, the disunity between experimental and behavioral economics in the United States has been amplified by opposing political ideologies. This is the reason why the distinction between experimental and behavioral economics is much more pronounced in the United States than it is in Europe. With their emphasis on the market as the only source of efficient allocations of resources and efficient production of goods, experimental economists such as Vernon Smith and Charles Plott are to be situated on the free-market end of the political spectrum, where the role of the government should be kept to an absolute minimum and where each individual has the

right and ability to take care of him- or herself. The repeated claim of Friedrich Hayek's intellectual heritage is another indication of how the experimental economists are politically positioned. Behavioral economists, on the other hand, favor a government that actively stimulates individuals to overcome their cognitive limitations, follow their own true preferences and thus advance their well-being. Behavioral economists favor a much more interventionist government, and implicitly a government that reduces the social inequality that results from economic and other structures that put some people at a disadvantage. Behavioral economists feel morally obliged to help individuals who, for example, are trapped by credit card companies into paying high interest rates on their loans, and to help individuals without considerable wealth to save for retirement.

Since the late 1970s experimental and behavioral economists have both relied on a branch of psychology called behavioral decision research to redefine economics in their own specific ways. The flow of theories, methods and experimental results from psychology to economics and their loss of psychological connotations illustrate the disunity of economics and psychology. So do the two very different ways in which these theories, methods and experimental results were picked up and used by experimental and behavioral economists to redefine economics. Economics and psychology are disunified cultures and behavioral economics and experimental economics are disunified sub-cultures. Their history can only be understood by recognizing them as such.

References

Non-published material

David Krantz, email to author, August 11, 2008

David Krantz – interview with the author, Columbia University, New York, June 20, 2008

David Krantz personal archive, Columbia University, New York

George Loewenstein, email to author, June 16, 2008

Gerd Gigerenzer, email to author, July 12, 2008

R. Duncan Luce's archive, Harvard University, Boston

Robyn Dawes – interview with the author, Carnegie Mellon University, Pittsburgh, June 23, 2008

Stephen Richards, email to author, December 4, 2008

Published material

Adair, J. G. (1984). "The Hawthorne Effect: A Reconsideration of the Methodological Artifact." Journal of Applied Psychology **69**(2): 334-345.

Allais, M. (1953). "Le comportement de l'homme Rationnel devant de le risque: critique des postulats et axioms de l'école américaine." Econometrica **21**: 503-546.

Allen, R. G. D. (1938). Mathematical Analysis for Economists. London, Macmillan.

Angner, E. and G. Loewenstein (forthcoming a). Behavioral economics. Philosophy of Economics. U. Maki.

Angner, E. and G. Loewenstein (forthcoming b). The Foundations of Behavioral Economics. Princeton, Princeton University Press.

Arrow, K. J. (1951). Social Choice and Individual Values. New York, Wiley.

Arrow, K. J. (1982). "Risk Perception in Psychology and Economics." Economic Inquiry **20**: 1-9.

Augier, M. and J. G. March, Eds. (2004). Models of a Man: Essays in Memory of Herbert A. Simon. Cambridge, MIT Press.

Barkow, J. H., L. Cosmides, et al. (1992). The Adapted Mind, Evolutionary Psychology and the Generation of Culture. Oxford, Oxford University Press.

Ben-Shakar, G., G. Bornstein, et al. (2007). "Reciprocity and Emotions in Bargaining: Using Physiological and Self-Report Measures." Journal of Economic Psychology **28**(3): 314-323.

- Bernoulli, D. (1954 (1738)). "Exposition of a New Theory on the Measurement of Risk." Econometrica **22**(1): 23-36.
- Bertrand, M., S. Mullainathan, et al. (forthcoming). "Behavioral Economics and Marketing in Aid of Decision-Making among the Poor." Journal of Public Policy and Marketing.
- Boumans, M. (forthcoming). Measurement in Economics. Handbook of Philosophy of Economics. U. Maki.
- Boumans, M. (2004). "Measurement outside the laboratory." Proceedings Philosophy of Science Assoc. 19th Biennial Meeting - PSA2004: PSA 2004 Contributed Papers.
- Bower, G. H. (1994). "A turning point in mathematical learning theory." Psychological Review **101**(2): 290-300.
- Bridgman, P. W. (1927). The Logic of Modern Physics. Toronto, Macmillan.
- Bulmer, M. (2001). History of Social Survey. International encyclopedia of the social and behavioral sciences. N. J. Smelser and P. B. Baltes. Oxford, Elsevier: 14469-14473.
- Camerer, C. (2003). Behavioral game theory: experiments in strategic interaction. Princeton, Princeton University Press.
- Camerer, C., L. Babcock, et al. (1997). "Labor supply of New York City cabdrivers: one day at a time." Quarterly Journal of Economics **112**(2): 407-441.
- Camerer, C., S. Issacharoff, et al. (2003). "Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism"." University of Pennsylvania Law Review **151**(1211-1254).
- Camerer, C., G. Loewenstein, et al. (2005). "Neuroeconomics: How Neuroscience Can Inform Economics." Journal of Economic Literature **43**: 9-64.
- Campbell, N. R. (1957). Foundations of Science: the philosophy of theory and experiment. New York, Dover Publications.
- Carnap, R. (1950). Logical Foundations of Probability. Chicago, University of Chicago Press.
- Chave, E. J. and L. L. Thurstone (1928). The Measurement of Attitude, A Psychophysical Method and Some Experiments with a Scale for Measuring Attitude toward the Church. Chicago, The University of Chicago Press.
- Chomsky, N. A. (1957). Syntactic Structures. The Hague, Mouton.
- Coombs, C. H. (1983). Psychology and Mathematics: an essay on theory. Ann Arbor, University of Michigan Press.

Coombs, C. H., R. M. Dawes, et al. (1970). Mathematical Psychology: an elementary introduction, Prentice-Hall.

Cosmides, L. and J. Tooby (1996). "Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty." Cognition **58**: 1-73.

Cremer, D. d., M. Zeelenberg, et al. (2006). Social Animals and Economic Beings: On unifying Social Psychology and Economics. Social Psychology and Economics. D. d. Cremer, M. Zeelenberg and J. K. Murnighan. London, Lawrence Erlbaum: 3-16.

Cronqvist, H. and R. Thaler (2004). "Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience." American Economic Review, papers and proceedings **94**(2): 424-428.

Damasio, A. (2003). Looking for Spinoza - Joy, Sorrow, and the Feeling Brain. New York, Harcourt Books.

Danziger, K. (1990). Constructing the Subject, historical origins of psychological research. New York, Cambridge University Press.

Danziger, K. (1997). Naming the Mind, how Psychology found its language. London, SAGE Publications.

Daston, L. (1983). Rational Individuals versus Laws of Society. Probability since 1800. M. Heidelberger, L. Krüger and R. Rheinwald. Bielefeld, Universität Bielefeld: 7-26.

Daston, L. (1988). Classical Probability in the Enlightenment. Princeton, Princeton University Press.

Daston, L. and P. Galison (2007). Objectivity. New York, Zone Books.

Davis, J. B. (2003). The Theory of the Individual in Economics. New York, Routledge.

de Finetti, B. (1937). "La prévision: ses lois logiques, ses sources subjective." Annales de l'Institut Henri Poincaré **7**: 1-68.

de Finetti, B. (1949). "Le vrai et le probable." Dialectica **3**: 78-93.

de Finetti, B. (1951). Recent suggestions for the reconciliation of theories of probability Proceedings of the Second [1950] Berkeley Symposium on Mathematical Statistics and Probability. J. Neyman. Berkeley, University of California Press: 217-226.

Debreu, G. (1954). Representation of preference ordering by a numerical function. Decision processes. R. M. Thrall, C. H. Coombs and R. L. Davis. New York, Wiley: 159-165.

- Debreu, G. (1958). "Stochastic choice and cardinal utility." Econometrica **26**: 440-444.
- Debreu, G. (1959a). "Cardinal utility for even-chance mixtures of pairs of sure prospects." Review of economic studies **71**: 174-177.
- Debreu, G. (1959b). Theory of Value: an axiomatic analysis of economic equilibrium. New York, Wiley.
- Debreu, G. (1960). Topological methods in cardinal utility theory. Mathematical methods in the social science, 1959. K. Arrow, S. Karlin and P. Suppes. Stanford, Stanford University Press: 16-26.
- Dimand, R. W. (2005). Experimental economic games: the early years. The Experiment in the History of Economics. P. Fontaine and R. Leonard. New York, Routledge: 5-24.
- Edwards, W. (1954). "The Theory of Decision Making." Psychological Bulletin **51**: 380-417.
- Edwards, W. (1961). "Behavioral decision theory." Annual Review of Psychology(12): 473-498.
- Edwards, W. and A. Tversky, Eds. (1967). Decision Making. Penguin Modern Psychology Readings. Baltimore, Penguin Books.
- Elster, J. and G. Loewenstein (1992). Utility from Memory and Anticipation. Choice over Time. G. Loewenstein and J. Elster. New York, Russell Sage Foundation: 213-234.
- Eriksson, L. and A. Hajek (forthcoming). "What Are Degrees of Belief?" Studia Logica.
- Estes, W. K. (1964). Probability learning. Categories of human learning. A. W. Melton. New York, Academic Press.
- Fama, E. (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work." Journal of Finance **25**(2): 383-417.
- Fechner, G. T. (1964/1860). Elemente der Psychophysik. Amsterdam, Bonset.
- Fishburn, P. C. (1964). Decision and Value Theory. New York, Wiley.
- Flom, M. C., F. W. Weymouth, et al. (1963). "Visual Resolution and Contour Interaction." Journal of the Optical Society of America **53**(9): 1026-1032.
- Frantilla, A. (1998). Social Science in the Public Interest: A fiftieth-year history of the Institute of Social Research. Ann Arbor, Bentley Historical Library, University of Michigan.

Frederick, S., G. Loewenstein, et al. (2002). "Time Discounting and Time Preference A Critical Review." Journal of Economic Literature **40**: 351-401.

Frederiksen, N. and H. Gulliksen, Eds. (1964). Contributions to Mathematical Psychology. London, Holt, Rinehart and Winston.

Friedman, M. and L. Savage (1948). "The utility analysis of choice involving risk." Journal of Political Economy(LVI 4): 279-304.

Friedman, M. and L. Savage (1952). "The expected-utility hypothesis and the measurability of utility." Journal of Political Economy(LX 6): 463-474.

Frijda, N. H. (1986). The Emotions. Cambridge, Cambridge University Press.

Fryback, D. (2005). "Ward Edwards: Father of Behavioral Decision Theory." Medical Decision Making **25**: 468-470.

Gabaix, X. and D. Laibson (2000). "A Boundedly Rational Decision Algorithm." The American Economic Review **90**(2): 433-438.

Galison, P. (1999). Trading Zone, Coordinating Action and Belief. the Science Studies reader. M. Biagioli. London Routledge: 137-160.

Galison, P. and D. J. Stump (1996). The Disunity of Science : Boundaries, Contexts, and Power. Stanford, Stanford University Press.

Gigerenzer, G. (1987a). The Probabilistic Revolution in Psychology - an Overview. The Probabilistic Revolution 2. L. Krüger, G. Gigerenzer and M. S. Morgan. Cambridge, MIT Press: 7-10.

Gigerenzer, G. (1987b). Probabilistic Thinking and the Fight against Subjectivity. The Probabilistic Revolution 2. L. Krüger, G. Gigerenzer and M. S. Morgan. Cambridge, MIT Press: 11-33.

Gigerenzer, G. (1987c). Survival of the Fittest Probabilist: Brunswik, Thurstone, and the Two Disciplines of Psychology. The Probabilistic Revolution 2. L. Krüger, G. Gigerenzer and M. S. Morgan. Cambridge, MIT Press: 49-72.

Gigerenzer, G. (1991). "How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases"." European Review of Social Psychology **2**: 83-115.

Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. Rationality, Psychological and Philosophical Perspectives. K. I. Manktelow and P. E. Over. London, Routledge: 284-313.

Gigerenzer, G. (1996). "On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky (1996)." Psychological Review **103**(3): 592-596.

- Gigerenzer, G. and D. J. Murray (1987). Cognition as Intuitive Statistics. Hillsdale, Lawrence Erlbaum.
- Gigerenzer, G. and R. Selten (2001). Bounded Rationality, The Adaptive Toolbox. Cambridge, MIT Press.
- Gigerenzer, G., P. M. Todd, et al., Eds. (1999). Simple Heuristics That Make Us Smart. Oxford, Oxford University Press.
- Green, D. M. and J. A. Swets (1966). Signal Detection Theory and Psychophysics. New York, Wiley.
- Grether, D. and C. R. Plott (1979). "Economic Theory of Choice and the Preference Reversal Phenomenon." The American Economic Review **69**(4): 623-638.
- Grether, D. M. and C. R. Plott (1982). "Economic Theory of Choice and the Preference Reversal Phenomenon: Reply." American Economic Review **72**(3): 575.
- Grubb, M. D. (2006). "Selling to Overconfident Consumers." Stanford Working Paper.
- Guilford, J. P. (1954). Psychometric Methods. New York, McGraw-Hill.
- Hacking, I. (1975). The emergence of probability : a philosophical study of early ideas about probability, induction and statistical inference. London, New York, Cambridge University Press.
- Hajek, A. (2007). Interpretations of Probability. The Stanford Encyclopedia of Philosophy (Winter 2007 Edition). E. N. Zalta.
- Harris, C. and D. Laibson (2001). "Dynamic Choices of Hyperbolic Consumers." Econometrica **69**(4): 935-957.
- Heidelberger, M. (1993). "Fechner's impact for measurement theory (comment on Murray, D.J 'A perspective for viewing the history of psychophysics')." Behavioral and Brain Sciences **16**: 146-148.
- Heidelberger, M. (2004). Nature from Within, Gustav Theodor Fechner and His Psychophysical Worldview. Pittsburgh, University of Pittsburgh Press.
- Heise, D. R. (1970). The semantic Differential and Attitude Research. Attitude Measurement. G. F. Summers. Chicago, Rand McNally: 235-253.
- Henrich, J. (2001). Group Report: What Is the Role of Culture in Bounded Rationality? Bounded Rationality, The Adaptive Toolbox. G. Gigerenzer and R. Selten. Cambridge USA, MIT Press.
- Henrich, J., R. Boyd, et al. (2001). "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies." The American Economic Review **91**(2): 73-78.

- Henrich, J., R. Boyd, et al. (2004). Foundations of human sociality : economic experiments and ethnographic evidence from fifteen small-scale societies Oxford, Oxford University Press.
- Hertwig, R. and G. Gigerenzer (1999). "The 'Conjunction Fallacy' Revisited: How Intelligent Inferences Look Like Reasoning Errors." Journal of Behavioral Decision Making **12**: 275-305.
- Hertwig, R. and A. Ortmann (2008). "Deception in Experiments: Revisiting the Arguments in Its Defense." Ethics and Behavior **18**(1): 59-92.
- Heukelom, F. (2005). "Gigerenzer the Decided, a tale of difficult distinctions." Tinbergen Institute Discussion Paper **2005-113/2**.
- Heukelom, F. (2007). "What Simon says." Tinbergen Institute discussion paper **2007-005/1**.
- Hollinger, D. A. (1989). Academic Culture at Michigan, 1938-1988: The Apotheosis of Pluralism. Intellectual History and Academic Culture at the University of Michigan: Fresh Explorations. M. A. Lourie. Ann Arbor, H. Rackham School of Graduate Studies: 58-101.
- House, J. S., F. T. Juster, et al., Eds. (2004). A Telescope on Society, Survey Research & Social Science at the University of Michigan & Beyond. Michigan, University of Michigan Press.
- Hyman, H. H. (1991). Taking Society's Measure. New York, Russell Sage.
- Innocenti, A. (2008). "How Can a Psychologist Inform Economics? The strange case of Sidney Siegel." DEPEID Working papers **8/2008**.
- Jallais, S. and P.-C. Pradier (2005). The Allais Paradox and its immediate consequences for expected utility theory. The Experiment in the History of Economics. P. Fontaine and R. Leonard. New York, Routledge.
- Jamisona, J., D. Karlaub, et al. (2008). "To Deceive or Not To Deceive: The effect of deception on behavior in future laboratory experiments." Journal of Economic Behavior & Organization **68**(3-4): 477-488.
- Jorland, G. (1987). The Saint Petersburg Paradox 1713-1937. The Probabilistic Revolution 1. L. Kruger, L. J. Daston and M. Heidelberger. Cambridge, MIT Press: 157-190.
- Jovanovic, F. (2008). "The Construction of the Canonical History of Financial Economics." History of Political Economy **40**(2): 213-242.
- Juster, F. T. (2004). The Behavioral Study of Economics. A Telescope on Society, Survey Research & Social Science at the University of Michigan & Beyond. J. S.

House, F. T. Juster, R. L. Kahn, H. Schuman and E. Singer. Ann Arbor, The University of Michigan Press.

Kahneman, D. (1963). "The semantic differential and the structure of inferences among attributes." American Journal of Psychology **76**: 554-567.

Kahneman, D. (1964). "Temporal summation in an acuity task at different energy levels - a study of the determinants of summation." Vision Research **4**: 557-566.

Kahneman, D. (1965a). "Control of spurious association and the reliability of the controlled variable." Psychological Bulletin **64**(5): 326-329.

Kahneman, D. (1965b). "Exposure duration and effective figure-ground contrast." Quarterly Journal of Experimental Psychology **17**: 308-314.

Kahneman, D. (1966a). "Time-intensity reciprocity in acuity as a function of luminance and figure-ground contrast." Vision Research **6**: 207-215.

Kahneman, D. (1966b). "Time-intensity reciprocity under various conditions of adaptation and backward masking." Journal of Experimental Psychology **71**: 543-549.

Kahneman, D. (1967). "An onset-onset law for one case of apparent motion and metacontrast." Perception & Psychophysics **2**: 577-584.

Kahneman, D. (1973). Attention and Effort. Englewoods Cliffs, Prentice Hall.

Kahneman, D. (2002). "Autobiography." from <http://nobelprize.org/economics/laureates/2002/kahneman-autobio.html>.

Kahneman, D. (2003). "Maps of Bounded Rationality: Psychology for Behavioral Economics." The American Economic Review **93**(5): 1449-1475.

Kahneman, D. and D. Beatty (1966). "Pupil Diameter and Load on Memory." Science **154**: 1583-1585.

Kahneman, D. and D. Beatty (1967). "Pupillary responses in a pitch-discrimination task." Perception and Psychophysics **2**: 101-105.

Kahneman, D., D. Beatty, et al. (1967). "Perceptual deficit during a mental task." Science **157**: 218-219.

Kahneman, D. and E. E. Ghiselli (1962). "Validity and Nonlinear Heteroscedastic Models." Personnel Psychology **15**: 1-12.

Kahneman, D., J. L. Knetsch, et al. (1990). "Experimental Tests of the Endowment Effect and the Coase Theorem." The Journal of Political Economy **98**(6): 1325-1348.

Kahneman, D. and J. Norman (1964). "The time-intensity relation in visual perception as a function of observer's task." Journal of Experimental Psychology **68**: 215-220.

- Kahneman, D., J. Norman, et al. (1967). "Critical duration for the resolution of form: centrally or peripherally determined." Journal of Experimental Psychology **73**(3): 323-327.
- Kahneman, D. and W. S. Peavler (1969). "Incentive effects and pupillary changes in association learning." Journal of Experimental Psychology **79**(2): 312-318.
- Kahneman, D., W. S. Peavler, et al. (1968). "Effect of verbalization and incentive on the pupil response to mental activity." Canadian Journal of Psychology **22**(3): 186-196.
- Kahneman, D., B. Tursky, et al. (1969). "Pupillary, heart rate, and skin resistance changes during a mental task." Journal of Experimental Psychology **79**(1): 164-167.
- Kahneman, D. and A. Tversky (1972). "Subjective Probability: A Judgment of Representativeness." Cognitive Psychology **3**: 430-454.
- Kahneman, D. and A. Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk." Econometrica **47**: 313-327.
- Kahneman, D. and A. Tversky (1996). "On the Reality of Cognitive Illusions." Psychological Review **103**(3): 582-591.
- Kahneman, D. and P. Wright (1971). "Changes of pupil size and rehearsal strategies in a short-term memory task." Quarterly Journal of Experimental Psychology **23**: 187-196.
- Kalish, G. K., J. W. Milnor, et al. (1954). Some Experimental N-person Games. New York, John Wiley & Sons.
- Keynes, J. M. (1921). Treatise on Probability. London, Macmillan & Co.
- Kiddler, L. M. (1981). Research Methods in Social Relations. New York, Holt, Rinehart & Winston.
- Klein, J. L. (2000). Economics for a Client: The Case of Statistical Quality Control and Sequential Analysis. Toward a History of Applied Economics. R. Backhouse and J. E. Biddle. Durham, Duke University Press. **32**: 27-69.
- Krantz, D. H., R. D. Luce, et al. (1971). Foundations of Measurement I. New York, Academic Press.
- Krüger, L., L. J. Daston, et al., Eds. (1987). The Probabilistic Revolution Volume 1: Ideas in History. Cambridge, MIT Press.
- Krüger, L., G. Gigerenzer, et al., Eds. (1987). The Probabilistic Revolution Volume 2: Ideas in Science. Cambridge, MIT Press.
- Laibson, D. (1997). "Golden Eggs and Hyperbolic Discounting." The Quarterly Journal of Economics **112**(2): 443-477.

- Laming, D. (1973). Mathematical Psychology. London, Academic Press.
- Lee, K. S. (2004). Rationality, Minds, and Machines in the Laboratory: A Thematic History of Vernon Smith's Experimental Economics. Notre Dame, University of Notre Dame. **PhD**: 300.
- Lewin, S. (1996). "Economics and Psychology: Lessons for Our Own Day From the Early Twentieth Century." Journal of Economic Literature **34**(3): 1293-1323.
- Loewenstein, G. (1992). The Fall and Rise of Psychological Explanations in the Economics of Intertemporal Choice. Choice over Time. G. Loewenstein and J. Elster. New York, Russell Sage Foundation: 3-34.
- Loewenstein, G. (1999). "Because It Is There: The Challenge of Mountaineering ... for Utility Theory." Kyklos **52**: 315-344.
- Loewenstein, G. (1999). "Experimental Economics from the vantage-point of Behavioural Economics." The Economic Journal **109**: F25-F34.
- Loewenstein, G. and J. Elster, Eds. (1992). Choice over Time. New York, Russell Sage Foundation.
- Loewenstein, G. and E. Haisley (forthcoming). The economist as therapist: Methodological issues raised by 'light' paternalism. Perspectives on the future of economics: Positive and normative foundations. A. Caplin and A. Schotter. Oxford, Oxford University Press. **1**.
- Loewenstein, G. and D. Prelec (1991). "Negative Time Preference." The American Economic Review **81**(2): 347-352.
- Loewenstein, G. and R. Thaler (1989). "Anomalies: Intertemporal choice." Journal of Economic Perspectives **3**: 181-193.
- Loewenstein, G., E. Weber, et al. (2001). "Risk as Feelings." Psychological Bulletin **127**: 267-286.
- Lopes, L. L. (1991). "The rhetoric of irrationality." Theory and Psychology **1**: 65-82.
- Luce, R. D., R. R. Bush, et al., Eds. (1963). The Handbook of Mathematical Psychology I. New York, John Wiley and Sons.
- Luce, R. D., R. R. Bush, et al., Eds. (1963). The Handbook of Mathematical Psychology II. New York, John Wiley and Sons.
- Luce, R. D., R. R. Bush, et al., Eds. (1965). The Handbook of Mathematical Psychology III. New York, John Wiley and Sons.
- Luce, R. D., D. H. Krantz, et al. (1990). Foundations of Measurement 3. Representation, axiomatization, and invariance. San Diego, Academic Press.

- Luce, R. D. and H. Raiffa (1957). Games and Decisions, Introduction and Critical Survey. New York, Dover Publications.
- Luce, R. D. and J. W. Tukey (1964). "Simultaneous Conjoint Measurement: A new type of fundamental measurement." Journal of Mathematical Psychology **1**(1): 1-27.
- Lunt, P. (1996). "Rethinking the relationship between economics and psychology." Journal of Economic Psychology **17**: 275-287.
- Maas, H. (2005). William Stanley Jevons and the making of modern economics. Cambridge, Cambridge University Press.
- MacHale, D. (1985). George Boole: His Life and Work. Dublin, Boole.
- MacKenzie, D. (2006). An Engine, Not a Camera, How Financial Models Shape Markets. London, MIT Press.
- Markowitz, H. M. (1952). "Portfolio Selection." Journal of Finance **7**(1): 77-91.
- Markowitz, H. M. (1959). Portfolio Selection: Efficient Diversification of Investments. New Jersey, John Wiley & Sons.
- Markowitz, H. M. (1965). "The Random Character of Stock Market Prices: A book Review." Journal of the American Statistical Association **60**: 309-381.
- McCabe, K. and V. L. Smith (2001). Goodwill Accounting and the Process of Exchange. Bounded Rationality, the Adaptive Toolbox. G. Gigerenzer and R. Selten. Cambridge, MIT Press: 319-342.
- McClure, S. M., D. Laibson, et al. (2004). "Separate Neural Systems Value Immediate and Delayed Monetary Rewards." Science **306**: 503-507.
- Michell, J. (1999). Measurement in Psychology. Cambridge, Cambridge University Press.
- Michell, J. (2007). Representational Theory of Measurement. Measurement in Economics: a handbook. M. Boumans. Amsterdam, Elsevier: 19-39.
- Miller, M. and F. Modigliani (1963). "Corporate Income Taxes and the Cost of Capital: A Correction." American Economic Review **53**(3): 433-443.
- Modigliani, F., A. B. Abel, et al. (1980). The Collected Papers of Franco Modigliani. Boston, MIT Press.
- Modigliani, F. and M. Miller (1958). "The Cost of Capital, Corporation Finance and the Theory of Investment." American Economic Review **48**(3): 261-297.
- Moscati, I. (2007). "Early Experiments in Consumer Demand Theory: 1930-1970." History of Political Economy **39**(3): 359-401.

- Mullainathan, S. (2007). Development Economics through the Lens of Psychology.
- Murnighan, J. K. and T. W. Ross (1999). "On the collaborative potential of psychology and economics." Journal of Economic Behavior and Organization **39**: 1-10.
- Murnighan, J. K. and A. E. Roth (2006). Some of the Ancient History of Experimental Economics and Social Psychology: Reminiscences and Analysis of a Fruitful Collaboration. Social Psychology and Economics. D. d. Cremer, M. Zeelenberg and J. K. Murnighan. London Lawrence Erlbaum: 321-333.
- Neumann, J. v. and O. Morgenstern (1944/2004). Theory of Games and Economic Behavior. Princeton, Princeton University Press.
- Ortmann, A. and R. Hertwig (2002). "The Costs of Deception: Evidence from Psychology." Experimental Economics **5**(2): 111-131.
- Osgood, C. E. (1953). Method and Theory in Experimental Psychology. New York, Oxford University Press.
- Peckham, H. H. (2005). The making of the University of Michigan 1917-1992. Ann Arbor, Bentley Historical Library, University of Michigan.
- Peterson, C. R. and L. R. Beach (1967). "Man as an Intuitive Statistician." Psychological Bulletin **88**(1): 29-46.
- Phillips, L. D. and D. Von Winterfeldt (2006). Reflections on the Contributions of Ward Edwards to Decision Analysis and Behavioral Research. LSEOR. London, LSE.
- Plott, C. R. and V. L. Smith, Eds. (2008). Handbook of Experimental Economics Results, Volume 1. Amsterdam, Elsevier.
- Poitras, G. and F. Jovanovic (2007). 20th Century Pioneers of Financial Economics. Cheltenham, Edward Elgar.
- Porter, T. (1994). "Making Things Quantitative." Science in Context **7**(3): 389-407.
- Porter, T. M. (1986). The Rise of Statistical Thinking 1820-1900. Princeton, Princeton University Press.
- Prelec, D. and G. Loewenstein (1997). "Beyond Time Discounting." Marketing Letters **8**(1): 97-108.
- Rabin, M. (1993). "Incorporating Fairness into Game theory and Economics." American Economic Review **83**(5): 1281-1302.
- Rabin, M. (1994). "Cognitive dissonance and social change." Journal of Economic Behavior and Organization **23**: 177-194.

- Rabin, M. (1998). "Psychology and Economics." Journal of Economic Literature **36**(1): 11-46.
- Ramsey, F. (1931/1950). The Foundations of Mathematics and other logical essays. London, Routledge.
- Rapoport, A. and A. Tversky (1970). "Choice Behavior in an Optimal Stopping Task." Organizational Behavior and Human Performance **5**: 105-120.
- Rosenbith, W. A. (1961). Sensory communication : contributions to the Symposium on principles of sensory communication. New York, Wiley.
- Ross, D. (2008). "Two Styles of Neuroeconomics." Economics and Philosophy **24**(373-383).
- Samuelson, P. A. (1947). Foundations of Economic Analysis. Cambridge, Harvard University Press.
- Santos, A. C. (2006). The Social Epistemology of Experimental Economics. Department of Philosophy. Rotterdam, Erasmus University Rotterdam. **PhD**.
- Savage, L. J. (1954). The Foundations of Statistics. New York, John Wiley & Sons.
- Schelling, T. (1957). "Bargaining, Communication, and Limited War." Journal of Conflict Resolution **1**: 19-36.
- Sedlmeier, P. and G. Gigerenzer (1997). "Intuitions About Sample Size: The Empirical Law of Large Numbers." Journal of Behavioral Decision Making **10**: 33-51.
- Sedlmeier, P. and G. Gigerenzer (2000). "Was Bernoulli Wrong? On Intuitions about Sample Size." Journal of Behavioral Decision Making **13**: 133-139.
- Sent, E.-M. (2004). "Behavioral Economics: How psychology made its (limited) way back into economics." History of Political Economy **36**(4): 735-760.
- Sent, E.-M. (2005). "Simplifying Herbert Simon." History of Political Economy **37**(2): 227-232.
- Shefrin, H. M. and R. Thaler (1988). "The Behavioral Life-Cycle Hypothesis." Economic Inquiry **26**(609-641).
- Simon, H. S. (1956). "Theories of Decision-Making in Economics and Behavioral Science." American Economic Review **49**(1): 253-283.
- Simon, H. S., Ed. (1957). Models of Man, Social and Rational - Mathematical Essays on Rational Human Behavior in a Social Setting. New York, John Wiley & Sons.
- Simon, H. S. (1959). "Theories of Decision-Making in Economics and Behavioral Sciences." American Economic Review **49**(1): 253-283.

- Simon, H. S. (1987). Behavioural Economics. The New Palgrave Fictionary of Economics. J. Eatwell, M. Milgate and O. Newman. London, Macmillan.
- Slovic, P. and S. Lichtenstein (1971). "Reversal of Preferences Between Bids and Choices in Gambling Decisions." Journal of Experimental Psychology **89**: 46-55.
- Slovic, P. and S. Lichtenstein (1973). "Response-Induced Reversals of Preferences in Gambling: An Extended Replication in Las Vegas." Journal of Experimental Psychology **101**: 16-20.
- Slovic, P. and S. Lichtenstein (1983). "Preference reversals: A broader perspective." American Economic Review **73**(4): 596-605.
- Smelser, N. J. (1986). The Ogburn Vision Fifty Years Later. Behavioral and Social Science: Fifty years of discovery. N. J. Smelser and D. R. Gerstein. Washington, National Academy Press: 21-35.
- Smith, V. L. (1959). "The Theory of Investment and Production." The Quarterly Journal of Economics **73**(1): 61-87.
- Smith, V. L. (1960). "Problems in Production-Investment Planning Over Time." International Economic Review **1**(3): 198-216.
- Smith, V. L. (1962). "An Experimental Study of Competitive Market Behavior." The Journal of Political Economy **70**(2): 111-137.
- Smith, V. L. (1965). "Experimental Auction Markets and the Walrasian Hypothesis." The Journal of Political Economy **73**(4): 387-393.
- Smith, V. L. (1967). "Experimental Studies of Discrimination Versus Competition in Sealed-Bid Auction Markets." The Journal of Business **40**(1): 56-84.
- Smith, V. L. (1968). "Economics of Production from Natural Resources." The American Economic Review **58**(3): 409-431.
- Smith, V. L. (1969). "Measuring Nonmonetary Utilities in Uncertain Choices: The Ellsberg Urn." The Quarterly Journal of Economics **83**(2): 324-329.
- Smith, V. L. (1974). "Economic Theory and Its Discontents." The American Economic Review **64**(2): 320-322.
- Smith, V. L. (1989). "Theory, Experiment and Economics." The Journal of Economic Perspectives **3**(1): 151-169.
- Smith, V. L. (2008). Rationality in Economics, Constructivist and Ecological Forms. Cambridge, Cambridge University Press.
- Stevens, S. S. (1939). "Psychology and the Science of Science." Psychological Bulletin **36**(4): 221-263.

Stevens, S. S., Ed. (1951). Handbook of Experimental Psychology. New York, John Wiley & Sons.

Stevens, S. S. (1951). Mathematics, Measurement, and Psychophysics. Handbook of Experimental Psychology. S. S. Stevens. New York, John Wiley & Sons: 1-50.

Suppes, P., D. H. Krantz, et al. (1989). Foundations of Measurement 2. Geometrical, threshold and probabilistic representations. San Diego, Academic Press.

Taussig, M. T. (1980). The devil and commodity fetishism in South America. Chapel Hill, The University of North Carolina press.

Teira, D. (2006). "On the Normative Dimension of the St. Petersburg Paradox." Studies in History and Philosophy of Science **37**: 210-223.

Thaler, R. (1980). "Toward a Positive Theory of Consumer Choice." Journal of Economic Behavior and Organization **1**: 7-59.

Thaler, R. (1987). The Psychology of Choice and the Assumptions of Economics. Laboratory Experiments in Economics: Six Points of View. A. Roth. Cambridge, Cambridge University Press.

Thaler, R., Ed. (1992). The winner's curse : paradoxes and anomalies of economic life. New York, Free Press.

Thaler, R., Ed. (1993). Advances in Behavioral Finance. New York, Russell Sage Foundation.

Thaler, R., Ed. (2005). Advances in Behavioral Finance. Princeton, Princeton University Press.

Thaler, R. H. (1991). Quasi Rational Economics. New York, Russell Sage Foundation.

Thaler, R. H. (2000). "From Homo Economicus to Homo Sapiens." Journal of Economic Perspectives **14**(1): 133-141.

Thaler, R. H. and S. Benartzi (2004). "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." Journal of Political Economy **112**: S164-187.

Thaler, R. H. and C. R. Sunstein (2003). "Libertarian Paternalism." The American Economic Review **93**(2): 175-179.

Thurstone, L. L. (1931). "The indifference function." Journal of Social Psychology **2**(2): 139-167.

Thurstone, L. L. (1927b). "A law of comparative judgment." Psychological Review **34**: 273-286.

Thurstone, L. L. (1959). The Measurement of Values Chicago, University of Chicago.

Thurstone, L. L. (1927c). "A mental unit of measurement." Psychological Review **34**: 415-423.

Thurstone, L. L. (1947). Multiple-factor Analysis : a development and expansion of The vectors of mind. Chicago, University of Chicago Press.

Thurstone, L. L. (1927a). "Psychophysical analysis." American Journal of Psychology **38**: 368-389.

Tversky, A. (1964). "On the Optimal Number of Alternatives at a Choice Point." Journal of Mathematical Psychology **1**: 386-391.

Tversky, A. (1967a). "Additivity, utility and subjective probability." Journal of Mathematical Psychology **4**: 175-202.

Tversky, A. (1967b). "A General Theory of Polynomial Conjoint Measurement." Journal of Mathematical Psychology **4**: 1-20.

Tversky, A. (1967c). "Utility theory and additivity analysis of risky choices." Journal of Experimental Psychology **75**(1): 27-36.

Tversky, A. (1969). "The intransitivity of preferences." Psychological Review **76**: 31-48.

Tversky, A. (1972). "Elimination by Aspects: A Theory of Choice." Psychological Review **79**: 218-299.

Tversky, A. (1991). Introduction. Frontiers of Mathematical Psychology, Essays in Honor of Clyde Coombs. D. R. Brown and J. E. K. Smith. New York, Springer-Verlag.

Tversky, A. and W. Edwards (1966). "Information versus reward in binary choices." Journal of Experimental Psychology **71**(5): 680-683.

Tversky, A. and D. Kahneman (1971). "Belief in the law of small numbers." Psychological Bulletin **76**: 105-110.

Tversky, A. and D. Kahneman (1973). "Availability: A heuristic for judging frequency and probability." Cognitive Psychology **5**: 207-232.

Tversky, A. and D. Kahneman (1974). "Judgment under Uncertainty: Heuristics and Biases." Science **185**: 1124-1131.

Tversky, A. and D. Kahneman (1980). Causal Schemas in Judgments under Uncertainty. Progress in Social Psychology. M. Fishbein. Hillsdale, Erlbaum: 49-72.

- Tversky, A. and D. Kahneman (1983). "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment." Psychological Review **91**: 293-315.
- Tversky, A. and D. H. Krantz (1970). "The Dimensional Representation and the Metric Structure of Similarity Data." Journal of Mathematical Psychology **7**: 572-596.
- Tversky, A. and J. E. Russo (1969). "Substitutability and Similarity in Binary Choices." Journal of Mathematical Psychology **6**: 1-12.
- Tversky, T. and D. Kahneman (1981). "The Framing of Decisions and the Psychology of Choice." Science **211**: 453-458.
- van Winden, F. (2007). "Affect and Fairness in Economics." Social Justice Research **20**(1): 32-52.
- von Plato, J. (1994). Creating Modern Probability. Cambridge, Cambridge University Press.
- Vromen, J. (2007). "Neuroeconomics as a Natural Extension of Bioeconomics: The Shifting Scope of Standard Economic Theory." Journal of Bioeconomics **9**(145-167).
- Wald, A. (1950). Statistical Decision Functions. New York, Wiley.
- Weintraub, E. R. (1991). Stabilizing dynamics, Constructing economic knowledge. Cambridge, Cambridge University Press.
- Weintraub, E. R., Ed. (1992). Toward a History of Game Theory. London, Duke University Press.
- Weintraub, E. R. and P. Mirowski (1994). "The Pure and the Applied: Bourbakism Comes to Mathematical Economics." Science in Context **7**(2): 245-272.

Samenvatting in het Nederlands

Dit proefschrift handelt over de invloed van twee psychologen, Daniel Kahneman en Amos Tversky, op de economische wetenschap in de afgelopen dertig jaar. Het is daarmee een historisch onderzoek naar de interactie tussen twee wetenschapsgebieden, de economie en de psychologie. Dat leidt meteen tot een eerste probleem, want hoe beschrijft men de interactie van twee culturen die verschillend naar de wereld, elkaar en de mens kijken? Om hier grip op te krijgen introduceer ik in het eerste, inleidende hoofdstuk Peter Galison's gebruik van het begrip 'disunity.' Disunity laat zich hier het beste vertalen als disharmonie. Economie en psychologie hebben veel met elkaar te maken en theorieën, instrumenten en experimentele resultaten reizen soms van de één naar de ander. Maar dat wil niet zeggen dat de twee één harmonieus geheel vormen. Het gebruik van de andere wetenschap dient altijd een specifiek doel binnen de eigen wetenschap en theorieën, instrumenten en experimentele resultaten die van de één naar de ander reizen verliezen in dit proces een gedeelte van de betekenis die ze in de wetenschap hadden waar ze vandaan komen, en krijgen nieuwe betekenis in de wetenschap van bestemming.

Met dit algemene raamwerk begint het tweede hoofdstuk met de mathematisch psychologen en keuzegedrag onderzoekers (behavioral decision researchers) aan de Universiteit van Michigan in de jaren 1950 en 1960. In deze periode was de Universiteit van Michigan het centrum van de Amerikaans psychologie. Het huisde onder andere het Institute for Social Research (ISR), Clyde Coombs' Michigan Mathematical Program, en Ward Edwards' Engineering Psychology Laboratory en het daaraan gekoppelde keuzegedrag onderzoek (behavioral decision research). Ik laat zien dat cruciaal voor de mathematische psychologie en keuzegedrag onderzoekers een tweezijdige interpretatie van psychologie was. Omdat de menselijke perceptie van de wereld afwijkt van de objectieve werkelijkheid is het noodzakelijk een theorie te ontwikkelen over hoe de mens die objectieve werkelijkheid waarneemt. Dit was de ene kant van psychologie. Maar aangezien meten uiteindelijk de keuze van het subject is tussen twee stimuli, was psychologie tegelijkertijd een onderzoek naar menselijk keuzegedrag. Met andere woorden, in de psychologie van Coombs' mathematische psychologie en Edwards' keuzegedrag onderzoek gebruikte men de mens als meetinstrument om menselijke keuzegedrag te meten.

Deze tweezijdige interpretatie van psychologie als het gebruik van een meetinstrument om datzelfde meetinstrument te onderzoeken werd problematisch toen duidelijk werd dat het meetinstrument zich niet gedroeg zoals het zich zou moeten gedragen. Dit was het probleem waar Tversky mee worstelde. Tversky moest kiezen tussen twee kwaden: *Of* zijn experimentele resultaten die lieten zien dat menselijk keuzegedrag vaak irrationeel is waren ongeldig, en de theorie van menselijk keuzegedrag toch in orde; *of* de experimentele resultaten waren geldig en de theorie fundamenteel onjuist. De laatste conclusie echter, zou eveneens meettheorie als basis voor de hele wetenschap ongeldig verklaren, een conclusie die Tversky niet zomaar bereid was te trekken. Kahneman bood een uitweg met de interpretatie dat er met de meettheorie en dus met de theorie van menselijk keuzegedrag niets mis was, maar dat dit een theorie was voor optimaal keuzegedrag waar daadwerkelijk, alledaags keuzegedrag op een voorspelbare manier van afweek. Aldus waren de experimentele resultaten geldig, terwijl tegelijkertijd de theorie voor meten en keuzegedrag behouden kon worden.

In het derde hoofdstuk behandel ik het werk van Kahneman en Tversky in de periode voor hun gezamenlijke onderzoek. Tversky deed onderzoek voor zijn PhD in de vroege jaren 1960 aan de Universiteit van Michigan. Zijn twee begeleiders waren Coombs en Edwards, en Tversky's onderzoek is een synthese van Coombs' mathematische psychologie en Edwards' keuzegedrag onderzoek. Tegen het einde van de jaren 1960 echter, worstelde Tversky steeds meer met de spanning tussen Leonard Savage's *a priori* keuzeaxioma's en het keuzegedrag van subjecten dat hij waarnam in zijn experimenten. Kahneman had een andere achtergrond. Sterk beïnvloed door zijn werk voor het Israëlische leger dat hij als student psychologie uitvoerde, concentreerde hij zich op de psychologie van cognitieve vergissingen. Kahneman liet zien dat alhoewel wij mensen denken dat we onze cognitie, onze hersenen, in het dagelijks leven goed gebruiken, dit allesbehalve het geval is en we systematisch cognitieve vergissingen maken.

De lange en succesvolle samenwerking tussen Kahneman en Tversky begon in 1969. De meest productieve periode waren de jaren 1970, die de basis legden voor hun roem in de psychologie, economie, en daarbuiten. In het vierde hoofdstuk zet ik Kahneman en Tversky's onderzoek uit de jaren 1970 uiteen, waarin ik laat zien hoe Kahneman's psychologie van cognitieve vergissingen een oplossing bood voor Tversky's worsteling met Savage's *a priori* axioma's en de experimentele

afwijkingen daarvan. Kahneman's oplossing was om de normatieve en descriptieve toepassing van de theorie rigoureus te scheiden. Daardoor konden Savage's *a priori* axioma's als normatieve regels voor keuzegedrag behouden blijven, terwijl tegelijkertijd de experimentele resultaten als descriptief bewijs konden dienen voor het feit dat menselijk keuzegedrag systematisch afwijkt van de norm. In 1979 culmineerde Kahneman en Tversky's onderzoek in "Prospect Theory: An Analysis of Decision under Risk," een artikel dat beschrijft hoe en waarom menselijk gedrag afwijkt van de norm. Prospect theory werd gepubliceerd in *Econometrica* en was expliciet bedoeld om ook de economen te overtuigen van de nieuwe psychologische benadering. In het artikel stelden Kahneman en Tversky dat economen en psychologen de verwachte nutstheorie (expected utility theory) op dezelfde manier gebruikten en betoogden daarmee impliciet dat economie en psychologie verbonden waren in één grote gedragswetenschap.

In hoofdstuk vijf laat ik zien hoe economen reageerden op prospect theory. Er vallen twee belangrijke reacties te onderscheiden, elk met hun eigen geschiedenis. Experimenteel economen zoals Vernon Smith bevestigden en accepteerden de experimentele resultaten, maar verwierpen elke oplossing, inclusief prospect theory, die de verwachte nutstheorie slechts aanpaste aan de nieuwe resultaten. Daarnaast concludeerden de experimenteel economen dat de resultaten van de psychologen het belang van de markt als een instituut dat over de tijd menselijk keuzegedrag naar een rationeel evenwicht stuurt alleen maar benadrukte. Financieel economen zoals Richard Thaler aan de ander kant accepteerden de experimentele resultaten van de psychologen eveneens, maar zagen ze als bewijs voor de waargenomen irrationaliteiten van financiële markten. Daarnaast haalden deze financieel economen Kahneman en Tversky en hun prospect theory binnen als de belangrijkste, zo niet de enige oplossing voor het theoretische gat dat was gevallen. Dit leidde tot de creatie van een nieuw gebied in de financiële economie, genaamd 'behavioral finance.' De reden voor het snelle succes van prospect theory was dat het de financieel economen een elegante weg uit de problemen bood. Het normatief-descriptief onderscheid zorgde ervoor dat de traditionele, neoklassieke modellen uit de economie gehandhaafd konden blijven als de normatieve theorie, terwijl tegelijkertijd een descriptief alternatief werd geboden dat maar weinig verschilde van de eerdere theorieën en dus makkelijk in de economie kon worden opgenomen.

In de late jaren 1980 en vroege jaren 1990 begonnen Thaler en zijn volgelingen hun benadering ook toe te passen op vraagstukken buiten de financiële economie. Het nieuwe, verbrede onderzoeksgebied groeide snel en werd in 1994 officieel gedragseconomie (behavioral economics) gedoopt. In het zesde hoofdstuk beschrijf ik de geschiedenis van de gedragseconomie in de jaren 1990 en 2000. Met behulp van de voorbeelden van onderzoek naar intertemporele keuze (intertemporal choice) en ontstane voorkeuren (emerging preferences) laat ik zien dat de gedragseconomen vele manieren onderzochten om nieuwe descriptieve theorieën te bouwen, maar uiteindelijk altijd trouw bleven aan het oorspronkelijk raamwerk van Kahneman en Tversky. Geleidelijk werden de predikten normatief en descriptief vervangen door volledige rationaliteit (full rationality) en beperkte rationaliteit (bounded rationality). Dit stelde gedragseconomen onder ander in staat om hun eigen visie op economisch beleidsadvies te ontwikkelen onder de kop van ‘paternalisme.’

Al deze verschillende ontwikkelingen droegen bij aan de gestage groei van de gedragseconomie tot een stabiel, helder omschreven en dominant economisch onderzoeksprogramma. Maar het maakte ook duidelijk hoe gedragseconomen hun programma onderscheidden van andere economische onderzoeksprogramma's en andere wetenschappelijk disciplines. Gedragseconomen begonnen met name hun programma te onderscheiden van de experimentele economie en van de psychologie. Dit lijkt enigszins schizofreen. Gedragseconomen definieerden hun programma als een verrijking van de economie met behulp van de psychologie, maar definieerden het tegelijkertijd nadrukkelijk als iets anders dan psychologie. En gedragseconomen maakten veelvuldig gebruik van experimenten en claimden het werk van Herbert Simon als belangrijke voorganger, maar onderscheidden gedragseconomie tegelijkertijd nadrukkelijk van experimentele economie, dat andere economische programma dat experimenten gebruikt en zichzelf als werkend in de Simon-traditie beschouwde.

Hoe moeten we deze schijnbare schizofrene in de gedragseconomie begrijpen? In het zevende en laatste hoofdstuk toon ik aan dat de geschiedenis beschreven in dit proefschrift laat zien hoe economen delen van de psychologie hebben gebruikt om de economische wetenschap te herdefiniëren. De reis van theorieën, methodes en experimentele resultaten van psychologie naar economie was nooit een neutraal proces dat theorieën, methodes en resultaten ongemoeid liet. In tegendeel, ze verloren een deel van hun psychologische connotatie en kregen er nieuwe economische

connotatie voor terug. Illustratief in dit verband is de geschiedenis van experimentele economie en gedragseconomie, die beiden zeer verschillende nieuwe connotaties gaven aan de psychologische theorieën, methoden, en resultaten. Experimentele economie en gedragseconomie gebruikten de psychologie om naar eigen inzicht, en op geheel eigen wijze de economische wetenschap te herdefiniëren. Daarmee laat dit proefschrift niet alleen zien dat theorieën, methoden, en experimentele resultaten die van de psychologie naar de economie reisden geen stabiele entiteiten waren, het benadrukt eveneens dat de definitie van wat de economische wetenschap is verre van constant is. Om die reden kan de geschiedenis van de interactie tussen economie en psychologie alleen begrepen worden door ze te zien als disharmonische culturen.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

405. M.Â. DOS REIS PORTELA, *Four essays on education, growth and labour economics.*
406. S.S. FICCO, *Essays on imperfect information-processing in economics.*
407. P.J.P.M. VERSIJP, *Advances in the use of stochastic dominance in asset pricing.*
408. M.R. WILDENBEEST, *Consumer search and oligopolistic pricing: A theoretical and empirical inquiry.*
409. E. GUSTAFSSON-WRIGHT, *Baring the threads: Social capital, vulnerability and the well-being of children in Guatemala.*
410. S. YERGOU-WORKU, *Marriage markets and fertility in South Africa with comparisons to Britain and Sweden.*
411. J.F. SLIJKERMAN, *Financial stability in the EU.*
412. W.A. VAN DEN BERG, *Private equity acquisitions.*
413. Y. CHENG, *Selected topics on nonparametric conditional quantiles and risk theory.*
414. M. DE POOTER, *Modeling and forecasting stock return volatility and the term structure of interest rates.*
415. F. RAVAZZOLO, *Forecasting financial time series using model averaging.*
416. M.J.E. KABKI, *Transnationalism, local development and social security: the functioning of support networks in rural Ghana.*
417. M. POPLAWSKI RIBEIRO, *Fiscal policy under rules and restrictions.*
418. S.W. BISSESSUR, *Earnings, quality and earnings management: the role of accounting accruals.*
419. L. RATNOVSKI, *A Random Walk Down the Lombard Street: Essays on Banking.*
420. R.P. NICOLAI, *Maintenance models for systems subject to measurable deterioration.*
421. R.K. ANDADARI, *Local clusters in global value chains, a case study of wood furniture clusters in Central Java (Indonesia).*
422. V.KARTSEVA, *Designing Controls for Network Organizations: A Value- Based Approach.*
423. J. ARTS, *Essays on New Product Adoption and Diffusion.*
424. A. BABUS, *Essays on Networks: Theory and Applications.*
425. M. VAN DER VOORT, *Modelling Credit Derivatives.*
426. G. GARITA, *Financial Market Liberalization and Economic Growth.*
427. E.BEKKERS, *Essays on Firm Heterogeneity and Quality in International Trade.*
428. H.LEAHU, *Measure-Valued Differentiation for Finite Products of Measures: Theory and Applications.*
429. G. BALTUSSEN, *New Insights into Behavioral Finance.*

430. W. VERMEULEN, *Essays on Housing Supply, Land Use Regulation and Regional Labour Markets.*
431. I.S. BUHAI, *Essays on Labour Markets: Worker-Firm Dynamics, Occupational Segregation and Workplace Conditions.*
432. C. ZHOU, *On Extreme Value Statistics.*
433. M. VAN DER WEL, *Riskfree Rate Dynamics: Information, Trading, and State Space Modeling.*
434. S.M.W. PHILIPPEN, *Come Close and Co-Create: Proximities in pharmaceutical innovation networks.*
435. A.V.P.B. MONTEIRO, *The Dynamics of Corporate Credit Risk: An Intensity-based Econometric Analysis.*
436. S.T. TRAUTMANN, *Uncertainty in Individual and Social Decisions: Theory and Experiments.*
437. R. LORD, *Efficient pricing algorithms for exotic derivatives.*
438. R.P. WOLTHOFF, *Essays on Simultaneous Search Equilibrium.*
439. Y.-Y. TSENG, *Valuation of travel time reliability in passenger transport.*
440. M.C. NON, *Essays on Consumer Search and Interlocking Directorates.*
441. M. DE HAAN, *Family Background and Children's Schooling Outcomes.*
442. T. ZAVADIL, *Dynamic Econometric Analysis of Insurance Markets with Imperfect Information.*
443. I.A. MAZZA, *Essays on endogenous economic policy.*
444. R. HAIJEMA, *Solving large structured Markov Decision Problems for perishable-inventory management and traffic control.*
445. A.S.K. WONG, *Derivatives in Dynamic Markets.*
446. R. SEGERS, *Advances in Monitoring the Economy.*
447. F.M. VIEIDER, *Social Influences on Individual Decision Making Processes.*
448. L. PAN, *Poverty, Risk and Insurance: Evidence from Ethiopia and Yemen.*
449. B. TIEBEN, *The concept of equilibrium in different economic traditions: A Historical Investigation.*
450. P. HEEMEIJER, *Expectation Formation in Dynamic Market Experiments.*
451. A.S. BOOIJ, *Essays on the Measurement Sensitivity of Risk Aversion and Causal Effects in Education.*
452. M.I. LÓPEZ YURDA, *Four Essays on Applied Microeconometrics.*
453. S. MEENTS, *The Influence of Sellers and the Intermediary on Buyers' Trust in C2C Electronic Marketplaces.*
454. S. VUJIĆ, *Econometric Studies to the Economic and Social Factors of Crime.*